

XX 01-MAY-1998; 98US-00071739.
PR
XX
XX (INSI-) INSIGHT STRATEGY & MARKETING LTD.
PA (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
PA (FRIE/) FRIEDMAN M M.
XX
XX Pecker I, Vlodavsky I, Friedman Y, Perets T;
XX MPI: 2000-052944/04.
DR N-PSDB; AA33290.
XX
XX Heparanase-specific molecular probes useful for diagnosis and treatment,
PT e.g. of tumors, and for targeted drug delivery.
XX
XX Example: Page 81-82; 90pp; English.
PS
XX The present invention describes heparanase-specific molecular probes,
CC useful for methods of detecting heparanase in situ. The probes and anti-
CC heparanase antibodies are used to detect or quantify the expression of
CC heparanase, for diagnosis and monitoring of diseases (especially
CC metastasis), for treatment of heparanase-associated diseases (e.g.
CC tumors, (adeno)carcinoma, squamous cell carcinoma, teratocarcinoma,
CC mesothelioma, melanoma, lymphoma or leukemia, a solid cancer (or its
CC metastases) derived from liver, prostate, bladder, breast, ovary, cervix,
CC colon, skin, intestine, stomach, uterus and pancreas, kidney disease,
CC diabetes and inflammation, haemorrhagic nephritis, nephrotic syndrome,
CC sepsis and inflammatory or autoimmune disease), for targeted drug
CC delivery (e.g. of anticancer agents) and as research reagents. The
CC present sequence represents human heparanase, which is used in the
CC exemplification of the present invention
XX
XX
SQ Sequence 543 AA;

Query Match 99.9%; Score 2838; DB 3; Length 543;

Best Local Similarity 99.8%; Pred. No. 5.3e-275;

Matches 542; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLTRKSPALPPMLMLLLGFLGFLSPGALPPAQAQDVVDLDPFTQEPHLVSPFLSVT 60
DB 1 MLTRKSPALPPMLMLLLGFLGFLSPGALPPAQAQDVVDLDPFTQEPHLVSPFLSVT 60
QY 61 IDANLATDPRFLILGSPFLRTIARGLSPAYLRFPGTKTDPLIFDPKKESTFEERSYQS 120
DB 61 IDANLATDPRFLILGSPFLRTIARGLSPAYLRFPGTKTDPLIFDPKKESTFEERSYQS 120
QY 121 QVNODICKYGSIPDPVEEKLRLWEPYQEQLLREHYQKFKNSTYSRSSVDVLYTFPANC 180
DB 121 QVNODICKYGSIPDPVEEKLRLWEPYQEQLLREHYQKFKNSTYSRSSVDVLYTFPANC 180
QY 181 GLDLIFGLNALRLRTADLQWSSNAQLLDYCSSKGYNISWEIGNEPNSFLKKADIFPINS 240
DB 181 GLDLIFGLNALRLRTADLQWSSNAQLLDYCSSKGYNISWEIGNEPNSFLKKADIFPINS 240
QY 241 QGGEDEFIQHLKLLRKSTFKNAKLYGPVQGPARRKTAQMLKSFLLKAGGEVIDSVTWHYYL 300
DB 241 QGGEDEYIQHLKLLRKSTFKNAKLYGPVQGPARRKTAQMLKSFLLKAGGEVIDSVTWHYYL 300
QY 301 NGRTATREDPLNDVLDIFISSVQKVFQVVESTPRGKTYWLGETSAYGGAPLLSDTPA 360
DB 301 NGRTATREDPLNDVLDIFISSVQKVFQVVESTPRGKTYWLGETSAYGGAPLLSDTPA 360
QY 361 AGFMWLDKGLSARMGIEVVMROVFFGAGNYHLVDENFPLPDYWLSTLFFKLVGTXVLM 420
DB 361 AGFMWLDKGLSARMGIEVVMROVFFGAGNYHLVDENFPLPDYWLSTLFFKLVGTXVLM 420
QY 421 ASVQSSKRRRLRVYHCTNTDNPXYKEGDLTYAINLHNVTKYRLPYPPSNKQVDKYL 480
DB 421 ASVQSSKRRRLRVYHCTNTDNPXYKEGDLTYAINLHNVTKYRLPYPPSNKQVDKYL 480
QY 481 RPLGPHGLSKSVQNLGLTKAVDDOTLPLMEKPLRRSSSLGLPAFSYSFVIRAKYA 540
DB 481 RPLGPHGLSKSVQNLGLTKAVDDOTLPLMEKPLRRSSSLGLPAFSYSFVIRAKYA 540

QY 541 ACT 543
DB 541 ACT 543

Search completed: March 23, 2005, 11:25:49
Job time : 114.5 secs

Db 301 NGRTATREDPLNPVDLIDFISVQKVFQVVESTPRGKVMWGTSAYGGAPLLSDTFA 360
 QY 361 AGFMWLDKLGISARNGIEVVMQVFFGAGNYHLVDENPDLPDYWLSTLFFKLVGTXYLM 420
 Db 361 AGFMWLDKLGISARNGIEVVMQVFFGAGNYHLVDENPDLPDYWLSTLFFKLVGTXYLM 420
 QY 421 ASVQGSKRKKLVYVLAHCTNTDNPVRYKEGDLTYALNLHVTKYLRLPYFSNKQVDKYL 480
 Db 421 ASVQGSKRKKLVYVLAHCTNTDNPVRYKEGDLTYALNLHVTKYLRLPYFSNKQVDKYL 480
 QY 481 RPLGPHGLLSKSVQNLGLTLKAVDDOTLPPLMEKPLRPSSSLGLPAFSYFFVIRNAKVA 540
 Db 481 RPLGPHGLLSKSVQNLGLTLKAVDDOTLPPLMEKPLRPSSSLGLPAFSYFFVIRNAKVA 540
 QY 541 ACT 543
 Db 541 ACT 543

RESULT 14

AAB08849 standard; protein; 543 AA.

AC AAB08849;
 DT 15-JAN-2001 (first entry)
 XX

DE Amino acid sequence of a human heparanase polypeptide.

XX Human; heparanase; gene therapy; tumour; inflammation; autoimmunity;
 KW heparin-binding growth factor; cytokine; neurodegenerative plaque;
 KW wound healing; infection; burn; angiogenesis; restenosis;
 KW atherosclerosis; inflammation; neurodegenerative disease;
 KW Gersmann-Strausler Syndrome; Creutzfeldt-Jakob disease.

XX Homo sapiens.

PN MO200052178-A1.

PD 08-SEP-2000.

PF 14-FEB-2000; 2000MO-US003542.

PR 01-MAR-1999; 99US-00258892.

PA (INSI-) INSIGHT STRATEGY & MARKETING LTD.
 PA (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
 PA (FRIE/) FRIEDMAN M M.

PI Pecker I, Vlodaevsky I, Feinstein E;

DR WPI; 2000-579289/54.

DR N-PSDB; AAA75051.

PT New polynucleotides encoding a polypeptide having heparanase activity,
 PT useful in wound healing and in gene therapy, particularly in treating
 PT tumor, inflammation, autoimmunity, neurodegenerative diseases.

PS Claim 22; Fig 1; 152pp; English.

XX The present sequence represents a human protein with heparanase catalytic
 CC activity. The heparanase (hpa) polynucleotide is useful in gene therapy,
 CC particularly in treating tumour, inflammation or autoimmunity.
 CC Particularly, the polynucleotide is useful in modulating the
 CC bioavailability of heparin-binding growth factors, cellular responses to
 CC heparin-binding growth factors (e.g. bFGF) and cytokines (e.g.
 CC interleukin (IL-8), cell interaction with plasma lipoproteins, cellular
 CC susceptibility to certain viral and some bacterial and protozoa
 CC infections, or disintegration of neurodegenerative plaques. The
 CC polynucleotide is also useful in wound healing (e.g. thermal, chemical or
 CC radiation burns), and in the treatment of angiogenesis, restenosis,
 CC atherosclerosis, inflammation, neurodegenerative diseases (Gersmann-
 CC Strausler Syndrome or Creutzfeldt-Jakob disease), and some viral,

CC bacterial or protozoa infections

XX Sequence 543 AA;

Query Match 99.9%; Score 2838; DB 3; Length 543;
 Best Local Similarity 99.8%; Pred. No. 5.3e-275;
 Matches 542; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLNSKRALPPMLMLLLGPGPLSPGALPPPAQADVDLDFFTOEPLHVSFSLSVT 60
 Db 1 MLNSKRALPPMLMLLLGPGPLSPGALPPPAQADVDLDFFTOEPLHVSFSLSVT 60
 QY 61 IDANLATDPPRFLILGSPKRLTLARGLSPALRFGCTTDPDLIDPKKESTFEERSYQWS 120
 Db 61 IDANLATDPPRFLILGSPKRLTLARGLSPALRFGCTTDPDLIDPKKESTFEERSYQWS 120
 QY 121 QVNODICKYGSIPPVVEKRLPLEMPYQQLLREHYQKKFNSTYSRSSVDVLTAFNCS 180
 Db 121 QVNODICKYGSIPPVVEKRLPLEMPYQQLLREHYQKKFNSTYSRSSVDVLTAFNCS 180
 QY 181 GDLIFGLNALRLRTADLQWSSNAQLLDYCSSKGYNISWEIGNEPNSFLKKADIFINGS 240
 Db 181 GDLIFGLNALRLRTADLQWSSNAQLLDYCSSKGYNISWEIGNEPNSFLKKADIFINGS 240
 QY 241 QLGEDFIQLHKLRLKSTFRNAKLYGPDYGOBRKTAKMLXSLFKAEGEVIDSVTHHYLL 300
 Db 241 QLGEDFIQLHKLRLKSTFRNAKLYGPDYGOBRKTAKMLXSLFKAEGEVIDSVTHHYLL 300
 QY 301 NGRTATREDPLNPVDLIDFISVQKVFQVVESTPRGKVMWGTSAYGGAPLLSDTFA 360
 Db 301 NGRTATREDPLNPVDLIDFISVQKVFQVVESTPRGKVMWGTSAYGGAPLLSDTFA 360
 QY 361 AGFMWLDKLGISARNGIEVVMQVFFGAGNYHLVDENPDLPDYWLSTLFFKLVGTXYLM 420
 Db 361 AGFMWLDKLGISARNGIEVVMQVFFGAGNYHLVDENPDLPDYWLSTLFFKLVGTXYLM 420
 QY 421 ASVQGSKRKKLVYVLAHCTNTDNPVRYKEGDLTYALNLHVTKYLRLPYFSNKQVDKYL 480
 Db 421 ASVQGSKRKKLVYVLAHCTNTDNPVRYKEGDLTYALNLHVTKYLRLPYFSNKQVDKYL 480
 QY 481 RPLGPHGLLSKSVQNLGLTLKAVDDOTLPPLMEKPLRPSSSLGLPAFSYFFVIRNAKVA 540
 Db 481 RPLGPHGLLSKSVQNLGLTLKAVDDOTLPPLMEKPLRPSSSLGLPAFSYFFVIRNAKVA 540
 QY 541 ACT 543
 Db 541 ACT 543

RESULT 15

AAV52990 standard; protein; 543 AA.

XX AAV52990;

DT 21-FEB-2000 (first entry)

DE Human heparanase protein sequence.

XX Human; heparanase; hpa; diagnosis; therapy; tumour; cytostatic;
 KW antidiabetic; immunomodulatory; anti-inflammatory; nephrotropic;
 KW metachelonia; adenocarcinoma; squamous cell carcinoma; teratocarcinoma;
 KW mesothelioma; melanoma; lymphoma; leukemia; cancer; sepsis; diabetes;
 KW inflammation; haemorrhagic nephritis; nephrotic syndrome;
 KW autoimmune disease; anticancer; kidney disease.

OS Homo sapiens.

PN WO9957153-A1.

XX 11-NOV-1999.

PF 29-APR-1999; 99MO-US009255.

disintegration of neurodegenerative plaques. Heparanase may be useful for conditions such as wound healing, angiogenesis, restenosis, atherosclerosis, inflammation, neurodegenerative diseases, and viral infections. Mammalian heparanase can be used to neutralize plasma heparin, and anti-heparanase antibodies may be applied for immunodetection and diagnosis of micrometastases, autoimmune lesions, and renal failure in biopsy specimens, plasma samples, and body fluids. The present sequence represents human heparanase

Sequence 543 AA:

Query Match 99.9%; Score 2838; DB 2; Length 543;
Best Local Similarity 99.8%; Pred. No. 5.3e-275;

Matches 542; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MLRSKRALPPMLMLLLGPGPLSPGALPRPAQADVVDDLFQOEPLHLVSPFLSVT 60
DB 1 MLRSKRALPPMLMLLLGPGPLSPGALPRPAQADVVDDLFQOEPLHLVSPFLSVT 60
QY 61 IDANLATPPRFLILGSPKRLTARGLSPAYLRFPGTKTDPLFPDKKESTFEERSYMQS 120
DB 61 IDANLATPPRFLILGSPKRLTARGLSPAYLRFPGTKTDPLFPDKKESTFEERSYMQS 120
QY 61 IDANLATPPRFLILGSPKRLTARGLSPAYLRFPGTKTDPLFPDKKESTFEERSYMQS 120
DB 61 IDANLATPPRFLILGSPKRLTARGLSPAYLRFPGTKTDPLFPDKKESTFEERSYMQS 120
QY 121 QVNODICKYGSIPDPVEEKLRLMPYQEQLLREHYQKKFNKSTYSRSSVDVLYTFANCS 180
DB 121 QVNODICKYGSIPDPVEEKLRLMPYQEQLLREHYQKKFNKSTYSRSSVDVLYTFANCS 180
QY 121 QVNODICKYGSIPDPVEEKLRLMPYQEQLLREHYQKKFNKSTYSRSSVDVLYTFANCS 180
DB 121 QVNODICKYGSIPDPVEEKLRLMPYQEQLLREHYQKKFNKSTYSRSSVDVLYTFANCS 180
QY 181 GLDLIFGNALLRTADLQWNSNAQLLDYCSSKQYNSWELGNPNSTLKKADIFINCS 240
DB 181 GLDLIFGNALLRTADLQWNSNAQLLDYCSSKQYNSWELGNPNSTLKKADIFINCS 240
QY 241 QLGEDFIQLHKLRLKSTFNKATLYGPDVGQPPRKATKMLKSLKAGGEYIDSVTMHHYLL 300
DB 241 QLGEDFIQLHKLRLKSTFNKATLYGPDVGQPPRKATKMLKSLKAGGEYIDSVTMHHYLL 300
QY 241 QLGEDFIQLHKLRLKSTFNKATLYGPDVGQPPRKATKMLKSLKAGGEYIDSVTMHHYLL 300
DB 241 QLGEDFIQLHKLRLKSTFNKATLYGPDVGQPPRKATKMLKSLKAGGEYIDSVTMHHYLL 300
QY 301 NGRTATREDPLNPVDLDFISSVQKVFQVVESTRPQKQWLGETSAYGAGAPLLSDTFA 360
DB 301 NGRTATREDPLNPVDLDFISSVQKVFQVVESTRPQKQWLGETSAYGAGAPLLSDTFA 360
QY 301 NGRTATREDPLNPVDLDFISSVQKVFQVVESTRPQKQWLGETSAYGAGAPLLSDTFA 360
DB 301 NGRTATREDPLNPVDLDFISSVQKVFQVVESTRPQKQWLGETSAYGAGAPLLSDTFA 360
QY 361 AGFWMLDLKGLSARNGIYVMQVFFGAGNYHLVDENPPLDVMYLSLFFKLVCTKYL 420
DB 361 AGFWMLDLKGLSARNGIYVMQVFFGAGNYHLVDENPPLDVMYLSLFFKLVCTKYL 420
QY 421 ASVQSSKRRLKRLVYHCTNTDNPRIKESGLTLAYANLHVTKYLLPYFPFNKQVDKYL 480
DB 421 ASVQSSKRRLKRLVYHCTNTDNPRIKESGLTLAYANLHVTKYLLPYFPFNKQVDKYL 480
QY 481 RPLGPHGLLSKSVQNLGLTLKAVDDQTLPLMEKPLRPSSSLGLPAFSYFVIRNAKYA 540
DB 481 RPLGPHGLLSKSVQNLGLTLKAVDDQTLPLMEKPLRPSSSLGLPAFSYFVIRNAKYA 540
QY 541 ACI 543
DB 541 ACI 543

```

RESULT 13
AA57590
AA57590 standard; protein; 543 AA.

AA57590;
02-MAR-2000 (first entry)

Human heparanase.

Human; heparanase; hpa; genetic modification; expression; anticancer;
angiogenesis; anti-angiogenic; antiproliferative; antiviral; antitumor;
anti-atherosclerotic; anti-inflammatory; antineurodegeneration;
heparan sulphate; heparin-binding growth factor; tumor angiogenesis;
metastasis; wound healing; restenosis; atherosclerosis; inflammation;
neurodegeneration; viral infection; cystic fibrosis; cancer; diagnosis;
micrometastasis; autoimmune lesion; kidney failure.

XX Homo sapiens.
OS
XX MO9957244-A1.
PN
XX 11-NOV-1999.
PD
XX 29-APR-1999; 99WO-US009256.
PF
XX 01-MAY-1998; 98US-00071618.
PR
XX 02-MAR-1999; 99US-00260038.
PA
XX (INSI-) INSIGHT STRATEGY & MARKETING LTD.
PA (FRIE/) FRIEDMAN M M.
XX Ben-Artzi H, Ayal-Hershenkovitz M, Yacoby-Zeevi O, Pecker I;
PI Peleg Y, Shlomi Y;
XX WPI, 2000-062144/05.
DR N-PSDB; AA239195.
XX
XX The present invention describes genetically modified cells (A) containing
CC a polynucleotide (I) that encodes a polypeptide with heparanase activity,
CC and express recombinant heparanase (II). Heparanase cleaves heparan
CC sulphate (HS) at specific intrachain sites, resulting in release of
CC heparin-binding growth factors, enzymes and proteins that are sequestered
CC by HS in basement membranes, extracellular matrix or cell surfaces. It
CC may also be implicated in tumor angiogenesis and metastases. (II) is
CC potentially useful in wound healing and for treating angiogenesis,
CC restenosis, atherosclerosis, inflammation, neurodegeneration, viral
CC infection and cystic fibrosis. It can also be used to neutralise heparin
CC (an alternative to protamine) and to screen for specific inhibitors
CC (potentially useful for treating cancer and metastases). Antibodies
CC raised against (II) are used for immunodetection and diagnosis of
CC micrometastases, autoimmune lesions and kidney failure. (A) provide (II)
CC in large quantities, in a form that is homogeneously processed and
CC activated/neutralised by a dedicated protease. The present sequence
CC represents human heparanase
XX
XX

Sequence 543 AA:

Query Match 99.9%; Score 2838; DB 3; Length 543;
Best Local Similarity 99.8%; Pred. No. 5.3e-275;

Matches 542; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MLRSKRALPPMLMLLLGPGPLSPGALPRPAQADVVDDLFQOEPLHLVSPFLSVT 60
DB 1 MLRSKRALPPMLMLLLGPGPLSPGALPRPAQADVVDDLFQOEPLHLVSPFLSVT 60
QY 61 IDANLATPPRFLILGSPKRLTARGLSPAYLRFPGTKTDPLFPDKKESTFEERSYMQS 120
DB 61 IDANLATPPRFLILGSPKRLTARGLSPAYLRFPGTKTDPLFPDKKESTFEERSYMQS 120
QY 61 IDANLATPPRFLILGSPKRLTARGLSPAYLRFPGTKTDPLFPDKKESTFEERSYMQS 120
DB 61 IDANLATPPRFLILGSPKRLTARGLSPAYLRFPGTKTDPLFPDKKESTFEERSYMQS 120
QY 121 QVNODICKYGSIPDPVEEKLRLMPYQEQLLREHYQKKFNKSTYSRSSVDVLYTFANCS 180
DB 121 QVNODICKYGSIPDPVEEKLRLMPYQEQLLREHYQKKFNKSTYSRSSVDVLYTFANCS 180
QY 121 QVNODICKYGSIPDPVEEKLRLMPYQEQLLREHYQKKFNKSTYSRSSVDVLYTFANCS 180
DB 121 QVNODICKYGSIPDPVEEKLRLMPYQEQLLREHYQKKFNKSTYSRSSVDVLYTFANCS 180
QY 181 GLDLIFGNALLRTADLQWNSNAQLLDYCSSKQYNSWELGNPNSTLKKADIFINCS 240
DB 181 GLDLIFGNALLRTADLQWNSNAQLLDYCSSKQYNSWELGNPNSTLKKADIFINCS 240
QY 181 GLDLIFGNALLRTADLQWNSNAQLLDYCSSKQYNSWELGNPNSTLKKADIFINCS 240
DB 181 GLDLIFGNALLRTADLQWNSNAQLLDYCSSKQYNSWELGNPNSTLKKADIFINCS 240
QY 241 QLGEDFIQLHKLRLKSTFNKATLYGPDVGQPPRKATKMLKSLKAGGEYIDSVTMHHYLL 300
DB 241 QLGEDFIQLHKLRLKSTFNKATLYGPDVGQPPRKATKMLKSLKAGGEYIDSVTMHHYLL 300
QY 241 QLGEDFIQLHKLRLKSTFNKATLYGPDVGQPPRKATKMLKSLKAGGEYIDSVTMHHYLL 300
DB 241 QLGEDFIQLHKLRLKSTFNKATLYGPDVGQPPRKATKMLKSLKAGGEYIDSVTMHHYLL 300
QY 301 NGRTATREDPLNPVDLDFISSVQKVFQVVESTRPQKQWLGETSAYGAGAPLLSDTFA 360
DB 301 NGRTATREDPLNPVDLDFISSVQKVFQVVESTRPQKQWLGETSAYGAGAPLLSDTFA 360

```

DT 20-MAR-2003 (revised)
 DT 14-OCT-1999 (first entry)
 XX A human protein with heparanase activity.
 DE Human; heparanase; heparan sulfate; trauma; autoimmune disease;
 KW skin disease; cardiovascular disease; nervous system disease;
 KW Alzheimer's disease; cancer; cancer metastasis; angiogenesis;
 KW inflammation; arthritis.
 XX Homo sapiens.
 OS WO9940207-A1.
 XX 12-AUG-1999.
 PD 05-FEB-1999; 99WO-EP000777.
 PF 09-FEB-1998; 98GB-00002725.
 PR (NOVS) NOVARTIS AG.
 PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
 XX Nakajima M, Toyoshima M;
 PI WPI; 1999-494300/41.
 XX N-PSDB; AAX86671.
 DR New heparanase polypeptide useful for treating autoimmune diseases, skin
 PT diseases, cardiovascular diseases and nervous system diseases including
 PT Alzheimer's disease.
 XX Claim 3; Page 29-31; 40pp; English.
 XX The present sequence represents a polypeptide with human heparanase
 CC biological activity. Antagonists and inhibitors of the protein prevent it
 CC from degrading the extracellular matrix and releasing heparan sulfate
 CC from the extracellular matrix surface. The heparanase protein or the anti-
 CC heparanase antibody are used in pharmaceutical compositions for treating
 CC warm blooded animals suffering from a disease resulting from shortage or
 CC lack of the heparanase protein, or from excessive activity or over-
 CC expression of the heparanase protein, respectively. The heparanase
 CC protein is used in treating diseases such as trauma, autoimmune disease,
 CC skin diseases, cardiovascular diseases and nervous system diseases
 CC including Alzheimer's disease resulting from shortage or lack of
 CC polypeptide. The anti-heparanase antibody is used in treating the
 CC diseases like cancer, cancer metastasis, angiogenesis and inflammation
 CC including arthritis resulting from excessive activity or over expression
 CC of heparanase protein. The anti-heparanase antibody can be used to detect
 CC the presence or absence of polypeptide and its concentration. (Updated on
 CC 20-MAR-2003 to correct PA field.)
 CC XX
 SQ Sequence 588 AA;
 Query Match 100.0%; Score 2841; DB 2; Length 588;
 Best Local Similarity 100.0%; Pred. No. 3e-275;
 Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 241 QLGEDFIQLHLKLRKSPFNKAKLYGPVCGPFRKTAKMLKSLKAGGEVIDSVTWHHYYL 300
 DB 286 QLGEDFIQLHLKLRKSTFPNKKLYGPDVCGPFRKTAKMLKSLKAGGEVIDSVTWHHYYL 345
 OY 301 NGRTATREDFLNDPVDLFISSVQKVFQVESSTRPGKRWLGTTSSAYGGGALLSDTFA 360
 DB 346 NGRTATREDFLNDPVDLFISSVQKVFQVESSTRPGKRWLGTTSSAYGGGALLSDTFA 405
 OY 361 AGFWMLDKLGLSARMGLEVVMRQVFGAGNYHLVDENFDDLPRYMLSLFKKLVTGTVLM 420
 DB 406 AGFWMLDKLGLSARMGLEVVMRQVFGAGNYHLVDENFDDLPRYMLSLFKKLVTGTVLM 465
 OY 421 ASVQSKRRKRLRYVLAHCTNTDNPYKGGDLTLVAINLHNTKYLRLPYPSNKQVDKTYL 480
 DB 466 ASVQSKRRKRLRYVLAHCTNTDNPYKGGDLTLVAINLHNTKYLRLPYPSNKQVDKTYL 525
 OY 481 RPLGPHGLSKSVOLNGLTLKMDVDDQTLPLMEKPLRPGSSGLPARFSFPVIRNAKVA 540
 DB 526 RPLGPHGLSKSVOLNGLTLKMDVDDQTLPLMEKPLRPGSSGLPARFSFPVIRNAKVA 585
 OY 541 ACI 543
 DB 586 ACI 588

RESULT 12
 ID AAY02345 standard; protein; 543 AA.
 XX AAY02345;
 AC AAY02345;
 XX 09-JUL-1999 (first entry)
 DT A human heparanase protein.
 DE Heparanase; hp; modulator; heparin-binding growth factor;
 KW cellular response; cytokine; cell interaction; plasma lipoprotein;
 KW cellular susceptibility; infection; disintegration;
 KW neurodegenerative plaque; wound healing; angiogenesis; restenosis;
 KW atherosclerosis; inflammation; neurodegenerative disease; neuritis;
 KW plasma heparin; micrometastasis; autoimmune lesion; renal failure.
 XX Homo sapiens.
 OS WO9911798-A1.
 XX 11-MAR-1999.
 PD 31-AUG-1998; 98WO-US017954.
 PF 02-SEP-1997; 97US-00922170.
 PR 02-JUL-1998; 98US-00109386.
 XX (INST-) INSTIGHT STRATEGY & MARKETING LTD.
 PA (HADA-) HADAST MEDICAL RES SERVICES & DEV.
 PA (FRIE/) FRIEDMAN M M.
 PI Pecker I, Vlodayevy I, Feinstein E;
 XX WPI; 1999-302255/25.
 DR N-PSDB; AAX35648.
 DB New human polynucleotide useful for treating angiogenesis, restenosis,
 PT and inflammation.
 PT Claim 6; Fig 1; 63pp; English.
 PS The specification describes a polypeptide having heparanase (hp)
 CC activity. The recombinant protein is used as a modulator of heparin-
 CC binding growth factors, cellular responses to heparin-binding growth
 CC factors and cytokines, cell interaction with plasma lipoproteins,
 CC cellular susceptibility to viral, protozoal and bacterial infections or


```

QY      481 RPLGPHGLSKSVOLNGTLTKMWDDOTLPPLMEKPLRPGSSGLPAPASYPFVIRNAKYA 540
Db      481 RPLGPHGLSKSVOLNGTLTKMWDDOTLPPLMEKPLRPGSSGLPAPASYPFVIRNAKYA 540
QY      541 ACI 543
Db      541 ACI 543

RESULT 9
ADR88210
ID      ADR88210 standard; protein; 543 AA.
AC      ADR88210;
XX      18-NOV-2004 (first entry)
DT      Human preproheparanase.
DE      Targeted drug delivery; inflammatory disorder; wound; scar; vasculopathy;
XX      autoimmune disorder; cancer; angiogenesis; metastatic disease;
XX      atherosclerosis; aneurysm; solid cancer; non-solid cancer;
XX      haematopoietic malignancy; lymphocytic leukaemia; myelogenous leukaemia;
XX      Hodgkin's disease; multiple myeloma; haemangiosarcoma; Kaposi's sarcoma;
XX      human; heparanase; enzyme.
OS      Homo sapiens.
FH      Key
FT      Location/Qualifiers
FT      Peptide
FT      1..35
FT      /label=Signal_peptide
FT      Protein
FT      36..543
FT      /label=Mature_heparanase
FT      Region
FT      36..109
FT      /note="8 kDa subunit of mature heparanase dimer"
FT      Domain
FT      89..107
FT      /note=Functional_peptide_epitope
FT      Region
FT      158..543
FT      /note="45 kDa subunit of mature heparanase dimer"
FT      Domain
FT      219..233
FT      /note=Functional_peptide_epitope
FT      Active-site
FT      225
FT      /note="Active site residue"
FT      Binding-site
FT      258..266
FT      /note="Putative heparin binding domain"
FT      Domain
FT      294..307
FT      /note=Functional_peptide_epitope
FT      Domain
FT      334..348
FT      /note=Functional_peptide_epitope
FT      Active-site
FT      343
FT      /note="Active site residue"
FT      Binding-site
FT      414..420
FT      /note="Putative heparin binding domain"
FT      Domain
FT      437..446
FT      /note=Functional_peptide_epitope
PN      US2004170631-A1.
XX      02-SEP-2004.
XX      28-NOV-2003; 2003US-00722502.
XX      02-SEP-1997; 97US-00922170.
XX      01-MAY-1998; 98US-00071739.
XX      01-NOV-1998; 98US-00186200.
XX      19-FEB-2003; 2003US-00368044.
XX      22-AUG-2003; 2003US-00645659.
XX      (YACO/) YACOBY-ZEEVI O.
XX      (PERE/) PERETZ T.
XX      (MIRON/) MIRON D.
XX      (SHLO/) SHLOMI Y.
XX      (PECK/) PECKER I.

```

```

PA      (AYAL/) AYAL-HERSHKOVITZ M.
PA      (FEIN/) FEINSTEIN E.
PA      (VGEH/) VAN GELDER J M.
PA      (VLGD/) VLADAVSKY I.
PA      (FRIE/) FRIEDMANN Y.
XX      Yacoby-Zeevi O, Peretz T, Miron D, Shlomi Y, Pecker I;
PI      Ayal-Hershkovitz M, Feinstein E, Van Gelder JM, Vladavsky I;
PI      Friedmann Y;
XX      MPI, 2004-625084/60.
DR      Claim 2; SEQ ID NO 4; 58pp; English.
XX      The invention relates to a method of targeted drug delivery to a tissue
XX      of a patient, the tissue expressing heparanase. The method comprises
XX      providing a complex of a drug directly or indirectly linked to an anti-
XX      heparanase antibody, and administering the complex to the patient. In the
XX      targeted drug delivery, the antibody comprises an antibody or its portion
XX      capable of specifically binding to at least one epitope of a heparanase
XX      protein. The composition and methods of the invention are useful for
XX      diagnosing, preventing or treating conditions associated with heparanase
XX      catalytic activity (e.g. an inflammatory disorder, wound, scar,
XX      vasculopathy, an autoimmune disorder, cancer, angiogenesis, cell
XX      proliferation, invasion of circulating tumour cells and metastatic
XX      disease), for purifying heparanase, or for developing drugs for those
XX      heparanase-associated conditions. The vasculopathy is atherosclerosis,
XX      restenosis or aneurysm. The cancerous condition is a solid cancer or a
XX      non-solid cancer. The non-solid cancer is a haematopoietic malignancy
XX      selected from acute lymphocytic leukaemia (ALL), acute myelogenous
XX      leukaemia (AML), myelodysplastic syndrome (MDS), mast cell leukaemia,
XX      Hodgkin's disease, non-Hodgkin's lymphomas, Burkitt's lymphoma and
XX      multiple myeloma. The solid cancer is selected from tumours in lip and
XX      oral cavity, pharynx, larynx, paranasal sinuses, major salivary glands,
XX      thyroid gland, oesophagus, stomach, small intestine, colon, colorectum,
XX      anal canal, liver, gallbladder, extrahepatic bile ducts, ampulla of
XX      Vater, exocrine pancreas, lung, pleural mesothelioma, soft tissue
XX      sarcoma, carcinoma and malignant melanoma of the skin, breast, vulva,
XX      vagina, cervix uteri, ovary, fallopian tube, gestational trophoblastic
XX      tumours, penis, prostate, testis, kidney, renal pelvis, ureter, urinary
XX      bladder, urethra, carcinoma of the eyelid, carcinoma of the conjunctiva,
XX      malignant melanoma of the conjunctiva, malignant melanoma of the uvea,
XX      retinoblastoma, carcinoma of the lacrimal gland, sarcoma of the orbit,
XX      brain, spinal cord, vascular system, haemangiosarcoma and Kaposi's
XX      sarcoma. The present sequence is human preproheparanase.
SQ      Sequence 543 AA:
QY      Query Match 100.0%; Score 2841; DB 8; Length 543;
Db      Best Local Similarity 100.0%; Pred. No. 2.7e-275;
QY      1 MLRSKRALPPMLLLGLPLSPGALPRPAQADVDLDFPTQPLHLVSPSFLSVT 60
Db      1 MLRSKRALPPMLLLGLPLSPGALPRPAQADVDLDFPTQPLHLVSPSFLSVT 60
QY      61 IDANLATDPRFLLILGSPKRLTLAGLSPAYLRFGGKTGTLFIDPKKESTFEERSYWG 120
Db      61 IDANLATDPRFLLILGSPKRLTLAGLSPAYLRFGGKTGTLFIDPKKESTFEERSYWG 120
QY      121 QVNODICKYGSIPDVDEKRLRLEWYQEQLLAEHYKRRKKNSTYSSSVLVYTFANCS 180
Db      121 QVNODICKYGSIPDVDEKRLRLEWYQEQLLAEHYKRRKKNSTYSSSVLVYTFANCS 180
QY      181 GDLIFGLNLRLPADLQNNSSNAQLLDYCSSKGYNISWELGNEPSPFLKKADIFNGS 240
Db      181 GDLIFGLNLRLPADLQNNSSNAQLLDYCSSKGYNISWELGNEPSPFLKKADIFNGS 240

```

PS Claim 9; SEQ ID NO 1296; 3699pp; English.

XX The invention relates to novel polynucleotide and polypeptides for
 CC treating psoriasis or a sequence having at least 80% identity to the
 CC above sequences. The nucleic acid is useful for preparing a composition
 CC for diagnosing or treating psoriasis in a mammal. This sequence
 CC corresponds to one of the polypeptides of the invention.

XX Sequence 543 AA;

Query Match 100.0%; Score 2841; DB 8; Length 543;

Best Local Similarity 100.0%; Pred. No. 2.7e-275;

Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLRSKPALPPMLLLGLPLSPGALPRPAQADVDVDFPQEPHLVSPFLSVT 60
 DB 1 MLRSKPALPPMLLLGLPLSPGALPRPAQADVDVDFPQEPHLVSPFLSVT 60
 QY 61 IDANLATDPRFLILGSPKLTARGLSPAYLRFGGTDTDFLIPDKKESTFEERSYWG 120
 DB 61 IDANLATDPRFLILGSPKLTARGLSPAYLRFGGTDTDFLIPDKKESTFEERSYWG 120
 QY 121 QVNODICKKGSIPPDVEEKLRLMPYQEQQLLREHYQKKFKKSTYSRSSVDVLYTFPANC 180
 DB 121 QVNODICKKGSIPPDVEEKLRLMPYQEQQLLREHYQKKFKKSTYSRSSVDVLYTFPANC 180
 QY 181 GDDLIFGLNALRLRTADLQWSSNAQLLDYCSSKGYNISWELGNEPNSFLKKADIFINCS 240
 DB 181 GDDLIFGLNALRLRTADLQWSSNAQLLDYCSSKGYNISWELGNEPNSFLKKADIFINCS 240
 QY 241 QLGEDFIQLHKLRLKSTFNKAKLYGPDVGQPRRTAKMLKSLFKAGGEVIDSVTMHHYVL 300
 DB 241 QLGEDFIQLHKLRLKSTFNKAKLYGPDVGQPRRTAKMLKSLFKAGGEVIDSVTMHHYVL 300
 QY 301 NGRTATREDFLNPVDLDFISSVQKVFQVVESTTRPGKKVWLGETSSAYGAGAPLSDTFA 360
 DB 301 NGRTATREDFLNPVDLDFISSVQKVFQVVESTTRPGKKVWLGETSSAYGAGAPLSDTFA 360
 QY 361 AGFMWLDKLGISARMGIEVVMRQVFFGAGNYHLVDENPDLDPDYMLSLFKVLGVTKVLM 420
 DB 361 AGFMWLDKLGISARMGIEVVMRQVFFGAGNYHLVDENPDLDPDYMLSLFKVLGVTKVLM 420
 QY 421 ASVQSSKRRKRLRVYLHCTNTDNPXYKEGDLTYAINLHNVTXYRLPYPFSSNKQVDKXYL 480
 DB 421 ASVQSSKRRKRLRVYLHCTNTDNPXYKEGDLTYAINLHNVTXYRLPYPFSSNKQVDKXYL 480
 QY 481 RPLGPHGLLSKSVQNLGLTKWVDDOTLPPLMEKPLRPSSSLGLPAFSSYFVIRNAKYA 540
 DB 481 RPLGPHGLLSKSVQNLGLTKWVDDOTLPPLMEKPLRPSSSLGLPAFSSYFVIRNAKYA 540
 QY 541 ACT 543
 DB 541 ACT 543

RESULT 8

ADQ80372

ID ADQ80372 standard; protein; 543 AA.

AC ADQ80372;

DT 21-OCT-2004 (first entry)

DE Heparinase protein.

KM cytosolic; epidermal growth factor receptor modulator; identification;

KW therapeutic response; cancer; EGFR; biomarker.

OS Homo sapiens.

PN WO2004063709-A2.

XX 29-JUL-2004.

XX 08-JAN-2004; 2004WO-US000368.

XX 08-JAN-2003; 2003US-0438735P.

PA (BRIM) BRISTOL-MYERS SQUIBB CO.

PI Amler LC, Januario T;

DR WPI: 2004-54414/52.

DR N-PSDB; ADQ80253.

PT Identifying a mammal that will respond therapeutically to a method of
 PT treating cancer comprises comparing the level of a biomarker in a mammal
 PT before and after exposure to an epidermal growth factor receptor (EGFR)
 PT modulator.

PS Disclosure; SEQ ID NO 144; 520pp; English.

XX The invention relates to a method of identifying a mammal that will
 CC respond therapeutically to a method of treating cancer by administering
 CC an epidermal growth factor receptor (EGFR) modulator by comparing the
 CC level of a biomarker in a mammal before and after exposure to an EGFR
 CC modulator. The method comprises: (a) measuring, in the mammal, the level
 CC of at least one biomarker identified in the specification; (b) exposing
 CC the mammal to the EGFR modulator; and (c) measuring in the mammal the
 CC level of the biomarker, where a difference in the level in step (c)
 CC compared to step (a) indicates that the mammal will respond
 CC therapeutically to the method of treating cancer. The method and
 CC biomarkers are useful for identifying a mammal that will respond
 CC therapeutically to a method of treating cancer by administering an
 CC epidermal growth factor receptor (EGFR) modulator. This sequence
 CC corresponds to one of the biomarkers whose levels of expression is
 CC measured in the method of the invention.

XX Sequence 543 AA;

Query Match 100.0%; Score 2841; DB 8; Length 543;

Best Local Similarity 100.0%; Pred. No. 2.7e-275;

Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLRSKPALPPMLLLGLPLSPGALPRPAQADVDVDFPQEPHLVSPFLSVT 60
 DB 1 MLRSKPALPPMLLLGLPLSPGALPRPAQADVDVDFPQEPHLVSPFLSVT 60
 QY 61 IDANLATDPRFLILGSPKLTARGLSPAYLRFGGTDTDFLIPDKKESTFEERSYWG 120
 DB 61 IDANLATDPRFLILGSPKLTARGLSPAYLRFGGTDTDFLIPDKKESTFEERSYWG 120
 QY 121 QVNODICKKGSIPPDVEEKLRLMPYQEQQLLREHYQKKFKKSTYSRSSVDVLYTFPANC 180
 DB 121 QVNODICKKGSIPPDVEEKLRLMPYQEQQLLREHYQKKFKKSTYSRSSVDVLYTFPANC 180
 QY 181 GDDLIFGLNALRLRTADLQWSSNAQLLDYCSSKGYNISWELGNEPNSFLKKADIFINCS 240
 DB 181 GDDLIFGLNALRLRTADLQWSSNAQLLDYCSSKGYNISWELGNEPNSFLKKADIFINCS 240
 QY 241 QLGEDFIQLHKLRLKSTFNKAKLYGPDVGQPRRTAKMLKSLFKAGGEVIDSVTMHHYVL 300
 DB 241 QLGEDFIQLHKLRLKSTFNKAKLYGPDVGQPRRTAKMLKSLFKAGGEVIDSVTMHHYVL 300
 QY 301 NGRTATREDFLNPVDLDFISSVQKVFQVVESTTRPGKKVWLGETSSAYGAGAPLSDTFA 360
 DB 301 NGRTATREDFLNPVDLDFISSVQKVFQVVESTTRPGKKVWLGETSSAYGAGAPLSDTFA 360
 QY 361 AGFMWLDKLGISARMGIEVVMRQVFFGAGNYHLVDENPDLDPDYMLSLFKVLGVTKVLM 420
 DB 361 AGFMWLDKLGISARMGIEVVMRQVFFGAGNYHLVDENPDLDPDYMLSLFKVLGVTKVLM 420
 QY 421 ASVQSSKRRKRLRVYLHCTNTDNPXYKEGDLTYAINLHNVTXYRLPYPFSSNKQVDKXYL 480
 DB 421 ASVQSSKRRKRLRVYLHCTNTDNPXYKEGDLTYAINLHNVTXYRLPYPFSSNKQVDKXYL 480

Db 241 QLGEDFIQLHLKLRKSTFNKAKLYGPDVGQPRRTAKMLKSLKAGGEVIDSVTHHHYLL 300
 QY 301 NGRATREDPLNPDVLDIFISSVQKVFQVVESTPRGKKVWLGETSSAYGGAPLLSDTFA 360
 Db 301 NGRATREDPLNPDVLDIFISSVQKVFQVVESTPRGKKVWLGETSSAYGGAPLLSDTFA 360
 QY 361 AGFMWLDKLGISARWGIEVVMROVFFGAGNYHLVDENFDPLPDYWLSTLFFKLVGTXYLM 420
 Db 361 AGFMWLDKLGISARWGIEVVMROVFFGAGNYHLVDENFDPLPDYWLSTLFFKLVGTXYLM 420
 QY 421 ASVQSKRRKLRVYLHCTNTDNPRIKEGDLTLVAIINLHVTKYLRPLPYFSNKQVDKXYLL 480
 Db 421 ASVQSKRRKLRVYLHCTNTDNPRIKEGDLTLVAIINLHVTKYLRPLPYFSNKQVDKXYLL 480
 QY 481 RPLGPHGLSKSVQVNLGLTLKAVDDQTLPLMEKPLRPGSSSLGLPAFSYFFVINAKYA 540
 Db 481 RPLGPHGLSKSVQVNLGLTLKAVDDQTLPLMEKPLRPGSSSLGLPAFSYFFVINAKYA 540
 QY 541 ACT 543
 Db 541 ACT 543

RESULT 6

ID ADN05074 standard; protein; 543 AA.

AC ADN05074;

DT 01-JUL-2004 (first entry)

DE Antipsoriatic protein sequence #716.

KW antipsoriatic; gene therapy; psoriasis; diagnosis.

OS Homo sapiens.

PN W02004028479-A2.

PD 08-APR-2004.

PF 25-SEP-2003; 2003WO-US030907.

PR 25-SEP-2002; 2002US-0414006P.

PA (GETH) GENENTECH INC.

PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;

PI Wu TD;

DR WPI: 2004-305105/28.

DR N-PSDB; ADN05073.

PT New PRO nucleic acid or polypeptide, useful for preparing a

PT pharmaceutical composition for diagnosing or treating psoriasis in a

PT mammal.

PS Claim 9; SEQ ID NO 1468; 3069pp; English.

CC The invention relates to novel polynucleotide and polypeptides for

CC treating psoriasis or a sequence having at least 80% identity to the

CC above sequences. The nucleic acid is useful for preparing a composition

CC for diagnosing or treating psoriasis in a mammal. This sequence

CC corresponds to one of the polypeptides of the invention.

XX Sequence 543 AA;

XX Query Match 100.0%; Score 2841; DB 8; Length 543;

XX Best Local Similarity 100.0%; Pred. No. 2,7e-275;

XX Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLNLSKPLPPPLMLLILGPIGLSPGLPRPAQADVVDLDFPTQGBLHVSPSLSVT 60

Db 1 MLNLSKPLPPPLMLLILGPIGLSPGLPRPAQADVVDLDFPTQGBLHVSPSLSVT 60
 QY 61 IDANLADPRRLILGSPKRLTLARGSPAYLRFGGTDTFLFDPKKESTFEERSYWOS 120
 Db 61 IDANLADPRRLILGSPKRLTLARGSPAYLRFGGTDTFLFDPKKESTFEERSYWOS 120
 QY 121 QVNODICKYGSIPPDVEBKRLLEMPYQDILLREHYQKKEFNSTYSRSSVDVLYTFANCS 180
 Db 121 QVNODICKYGSIPPDVEBKRLLEMPYQDILLREHYQKKEFNSTYSRSSVDVLYTFANCS 180
 QY 181 GLDLIFGINALLRTADLOMNSNNAQLLDYCSSKGVNISMELGNEPNSFLKADIFINGS 240
 Db 181 GLDLIFGINALLRTADLOMNSNNAQLLDYCSSKGVNISMELGNEPNSFLKADIFINGS 240
 QY 241 QLGEDFIQLHLKLRKSTFNKAKLYGPDVGQPRRTAKMLKSLKAGGEVIDSVTHHHYLL 300
 Db 241 QLGEDFIQLHLKLRKSTFNKAKLYGPDVGQPRRTAKMLKSLKAGGEVIDSVTHHHYLL 300
 QY 301 NGRATREDPLNPDVLDIFISSVQKVFQVVESTPRGKKVWLGETSSAYGGAPLLSDTFA 360
 Db 301 NGRATREDPLNPDVLDIFISSVQKVFQVVESTPRGKKVWLGETSSAYGGAPLLSDTFA 360
 QY 361 AGFMWLDKLGISARWGIEVVMROVFFGAGNYHLVDENFDPLPDYWLSTLFFKLVGTXYLM 420
 Db 361 AGFMWLDKLGISARWGIEVVMROVFFGAGNYHLVDENFDPLPDYWLSTLFFKLVGTXYLM 420
 QY 421 ASVQSKRRKLRVYLHCTNTDNPRIKEGDLTLVAIINLHVTKYLRPLPYFSNKQVDKXYLL 480
 Db 421 ASVQSKRRKLRVYLHCTNTDNPRIKEGDLTLVAIINLHVTKYLRPLPYFSNKQVDKXYLL 480
 QY 481 RPLGPHGLSKSVQVNLGLTLKAVDDQTLPLMEKPLRPGSSSLGLPAFSYFFVINAKYA 540
 Db 481 RPLGPHGLSKSVQVNLGLTLKAVDDQTLPLMEKPLRPGSSSLGLPAFSYFFVINAKYA 540
 QY 541 ACT 543
 Db 541 ACT 543

RESULT 7

ID ADN04902 standard; protein; 543 AA.

AC ADN04902;

DT 01-JUL-2004 (first entry)

DE Antipsoriatic protein sequence #631.

KW antipsoriatic; gene therapy; psoriasis; diagnosis.

OS Homo sapiens.

PN W02004028479-A2.

PD 08-APR-2004.

PF 25-SEP-2003; 2003WO-US030907.

PR 25-SEP-2002; 2002US-0414006P.

PA (GETH) GENENTECH INC.

PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;

PI Wu TD;

DR WPI: 2004-305105/28.

DR N-PSDB; ADN04901.

PT New PRO nucleic acid or polypeptide, useful for preparing a

PT pharmaceutical composition for diagnosing or treating psoriasis in a

PT mammal.

DB 541 ACT 543

RESULT 3
ADD18950
ID ADD18950 standard; protein; 543 AA.

AC ADD18950;
XX
XX 15-JAN-2004 (first entry)
XX

DE Human disease related protein SeqId439.
XX
XX human; disease state; cytostatic; antiinflammatory; ophthalmological;
XX antiarteriosclerotic; vulnery; gene therapy;
XX hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;
XX inflammation; erythropoiesis; glycolysis; gluconeogenesis;
XX glucose transportation; catecholamine synthesis; iron transport;
XX nitric oxide synthesis; cancer; ischemic condition; reperfusion injury;
XX retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
XX inflammatory condition; wound healing.

OS Homo sapiens.
XX
XX WO2003018621-A2.
XX
XX 06-MAR-2003.
XX
XX 23-AUG-2002; 2002WO-GB003892.
XX
XX 23-AUG-2001; 2001GB-00020558.
XX
XX 05-OCT-2001; 2001GB-00024037.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX
XX Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
XX
XX WPI; 2003-290046/28.
XX
XX N-PSDB; ADD18951.
XX
XX
XX New substantially purified polypeptide, useful for diagnosing or treating
XX a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
XX injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
XX wound healing.

PT Claim 25; SEQ ID NO 439; 424bp; English.

XX This invention relates to novel human genes and gene product which are
XX implicated in certain disease states. Compounds which modulate the
XX proteins of the invention may have cytostatic, antiinflammatory,
XX ophthalmological, antiarteriosclerotic or vulnery activities. The
XX sequences of the invention may be useful for gene therapy. The invention
XX may be useful for diagnosing or treating a hypoxia-regulated condition,
XX such as tumorigenesis, angiogenesis, apoptosis, inflammation,
XX erythropoiesis, or the biological response to hypoxia conditions
XX including processes such as glycolysis, gluconeogenesis, glucose
XX transportation, catecholamine synthesis, iron transport or nitric oxide
XX synthesis. The disease includes cancer, ischemic conditions, reperfusion
XX injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
XX inflammatory conditions or wound healing. The present sequence is that of
XX a disease related protein of the invention.

XX
XX
XX Sequence 543 AA;
XX

Query Match 100.0%; Score 2841; DB 7; Length 543;
Best Local Similarity 100.0%; Pred. No. 2.7e-275;
Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLRSKRALPPPLMLLLGLSPGLSPGALPRPAQADVDLDLDFTEOPLHLVSPSLSYT 60
DB 1 MLRSKRALPPPLMLLLGLSPGLSPGALPRPAQADVDLDLDFTEOPLHLVSPSLSYT 60
QY 61 IDANLATDPRFLILGSPKRLTLAGLSPAYLRFGGTKTDFLIIPDKKSTFEERSYWGOS 120

DB 61 IDANLATDPRFLILGSPKRLTLAGLSPAYLRFGGTKTDFLIIPDKKSTFEERSYWGOS 120

QY 121 QVNODICKYGSIPDPVEEKLRLLEMPYQEQLLREHYQKKFKNSTYSSSSVDVLYTFRANGS 180
DB 121 QVNQDICKYGSIPDPVEEKLRLLEMPYQEQLLREHYQKKFKNSTYSSSSVDVLYTFRANGS 180

QY 181 GDLDFGLNALIRTDALQWSSNAQLLDYCSSKGYNISWEIGNEPNSFLKKADIFINGG 240
DB 181 GDLDFGLNALIRTDALQWSSNAQLLDYCSSKGYNISWEIGNEPNSFLKKADIFINGG 240

QY 241 QGGEDEFOIHLKLRSTFKNAKLYGPDVGQPPRKTKMKLSFLKAGGEVIDSYTMHHYYL 300
DB 241 QGGEDEFOIHLKLRSTFKNAKLYGPDVGQPPRKTKMKLSFLKAGGEVIDSYTMHHYYL 300

QY 301 NGRTATREDFLNPVDVLDIFISSVQKVFQVESTRPCKKWLGETSSAYGGAPLSDTFA 360
DB 301 NGRTATREDFLNPVDVLDIFISSVQKVFQVESTRPCKKWLGETSSAYGGAPLSDTFA 360

QY 361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENPDLPDYWLSLFPKKLVGTRVLM 420
DB 361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENPDLPDYWLSLFPKKLVGTRVLM 420

QY 421 ASVQSKRRKLRVYLHCTNTDNPYKEGDLTYALNLHVTXYLRPYPSNKQVDKYL 480
DB 421 ASVQSKRRKLRVYLHCTNTDNPYKEGDLTYALNLHVTXYLRPYPSNKQVDKYL 480

QY 481 RPLGPHGLSKSVQNLGLTLKMWDDQTLPLMEKPLRPSSSGLPAPFSYFFVIRNAKVA 540
DB 481 RPLGPHGLSKSVQNLGLTLKMWDDQTLPLMEKPLRPSSSGLPAPFSYFFVIRNAKVA 540

QY 541 ACT 543
DB 541 ACT 543

RESULT 4
ADK52086
ID ADK52086 standard; protein; 543 AA.
XX
XX
XX ADK52086;
XX
XX 20-MAY-2004 (first entry)
XX
XX
XX Human atopic dermatitis/psoriasis-associated protein #1.
XX
XX Human atopic dermatitis; psoriasis; dermatological; anti-inflammatory;
XX antipsoriatic; rash.
XX
XX Homo sapiens.
XX
XX WO2004016785-A1.
XX
XX 26-FEB-2004.
XX
XX 06-AUG-2003; 2003WO-JP009999.
XX
XX 06-AUG-2002; 2002JP-00229319.
XX
XX 14-MAY-2003; 2003JP-00136544.
XX
XX (GENO-) GENOX RES INC.
XX (UYJU-) UNIV JUNTENDO.
XX
XX Itoh M, Ogawa K, Shinagawa A, Sudo H, Ogawa H, Ra C;
XX Mitsuishi K;
XX WPI; 2004-214514/20.
XX N-PSDB; ADK51968.
XX
XX Detecting atopic dermatitis or psoriasis comprises assaying levels of a
XX expression of an indicator gene at a rash site and non-rash site of a
XX person with atopic dermatitis or psoriasis.

CC above can be diagnosed using specific antibodies, and also using primers
CC and probes specific for the heparanase polynucleotides. Other uses of the
CC heparanases include sequencing sulfated molecules such as HSPG. The
CC present sequence represents a human heparanase

XX Sequence 543 AA;

Query Match 100.0%; Score 2841; DB 2; Length 543;
Best Local Similarity 100.0%; Pred. No. 2.7e-275; Indels 0; Gaps 0;
Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLRSKPALPPMLLLGLPLSPGALPRPAQADVVDLDFTEPHLHVSPSFLSVT 60
DB 1 MLRSKPALPPMLLLGLPLSPGALPRPAQADVVDLDFTEPHLHVSPSFLSVT 60
QY 61 IDANLATDRPFLILGSPKRLTARGLSPAYLRFGGTDTDFLIPDKKSTFEERSYWS 120
DB 61 IDANLATDRPFLILGSPKRLTARGLSPAYLRFGGTDTDFLIPDKKSTFEERSYWS 120
QY 121 QVNODICKYGIIPDVEEKLRLMPYQEQDLLREHYQKKFKNSTYSRSSVDVLYTFPANC 180
DB 121 QVNODICKYGIIPDVEEKLRLMPYQEQDLLREHYQKKFKNSTYSRSSVDVLYTFPANC 180
QY 181 GDLIFGLNALIRTDALQWSSNAQLLDYCSSKGYNISWELGNEBNSFLKKADIFINGS 240
DB 181 GDLIFGLNALIRTDALQWSSNAQLLDYCSSKGYNISWELGNEBNSFLKKADIFINGS 240
QY 241 QGGEPIQHLKLRSTFNKALYGPVQOPRRKTAKMKSFLKAGGEVIDSTVTHHYYL 300
DB 241 QGGEPIQHLKLRSTFNKALYGPVQOPRRKTAKMKSFLKAGGEVIDSTVTHHYYL 300
QY 301 NGRTATREDPLNPVDLDFISSVQKVFQVVESTPGKTYMLGETSSAYGGAPLISDTFA 360
DB 301 NGRTATREDPLNPVDLDFISSVQKVFQVVESTPGKTYMLGETSSAYGGAPLISDTFA 360
QY 361 AGFMWLDKLGISARMGIEVVMQVFFGAGNYHLVDENFDPLPDYWLSTLLFKKLVTGTYLM 420
DB 361 AGFMWLDKLGISARMGIEVVMQVFFGAGNYHLVDENFDPLPDYWLSTLLFKKLVTGTYLM 420
QY 421 ASVQSSKRRKRLRVYLHCTNDNPRYKGGDLTYAINLHVNTKYLRPPFSNKQVDKYL 480
DB 421 ASVQSSKRRKRLRVYLHCTNDNPRYKGGDLTYAINLHVNTKYLRPPFSNKQVDKYL 480
QY 481 RPLGPHGLSKSVQNLGLTLKKVDDQTLPLMEKPLRPGSSISGLPAFSYSPFVIRNAKYA 540
DB 481 RPLGPHGLSKSVQNLGLTLKKVDDQTLPLMEKPLRPGSSISGLPAFSYSPFVIRNAKYA 540
QY 541 ACT 543
DB 541 ACT 543

RESULT 2

AAB86206
ID AAB86206 standard; protein; 543 AA.

AC AAB86206;

XX 24-AUG-2001 (first entry)

XX Human heparanase inhibitor protein.

XX Heparanase; inhibitor; cardiac insufficiency; cardiatic; nephrotropic;

KM hepatocytic; veterinary medicine; congestive heart failure; dyspnoea;

KM primary cardiomyopathy; peripheral odema; pulmonary congestion;

XX hepatic congestion; hydrothorax; ascite; nocturia; human.

OS Homo sapiens.

XX DE19955803-A1.

XX 23-MAY-2001.

PF 19-NOV-1999; 99DE-01055803.
XX
PR 19-NOV-1999; 99DE-01055803.
XX

PA (KNOL) KNOL AG.

PI Herr D, Hahn A, Laux V;

XX WPI: 2001-368371/39.

DR N-PSDB; AAH20940.

PT Treatment or prevention of cardiac insufficiency and related conditions,
PT e.g. pulmonary congestion and dyspnoea, comprises administration of
PT heparanase inhibitor.

XX Disclosure; Page 11-13; 16pp; German.

CC This invention describes a novel heparanase inhibitor which can be used
CC for the treatment or prevention of cardiac insufficiency and associated
CC indications, symptoms and/or malfunctions. The heparanase inhibitor of
CC the invention has cardiatic, nephrotropic and hepatotropic activity. The
CC products of the invention can be used in human and veterinary medicine.
CC For the treatment or prevention of congestive heart failure e.g. primary
CC cardiomyopathy. Associated conditions treated or prevented with the
CC inhibitor are especially peripheral odemas, pulmonary and hepatic
CC congestion, dyspnoea, hydrothorax and ascites. Renal problems, e.g.
CC nocturia can also be treated. This sequence represents the human
CC heparanase protein described in the method of the invention

XX Sequence 543 AA;

Query Match 100.0%; Score 2841; DB 4; Length 543;
Best Local Similarity 100.0%; Pred. No. 2.7e-275; Indels 0; Gaps 0;
Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLRSKPALPPMLLLGLPLSPGALPRPAQADVVDLDFTEPHLHVSPSFLSVT 60
DB 1 MLRSKPALPPMLLLGLPLSPGALPRPAQADVVDLDFTEPHLHVSPSFLSVT 60
QY 61 IDANLATDRPFLILGSPKRLTARGLSPAYLRFGGTDTDFLIPDKKSTFEERSYWS 120
DB 61 IDANLATDRPFLILGSPKRLTARGLSPAYLRFGGTDTDFLIPDKKSTFEERSYWS 120
QY 121 QVNODICKYGIIPDVEEKLRLMPYQEQDLLREHYQKKFKNSTYSRSSVDVLYTFPANC 180
DB 121 QVNODICKYGIIPDVEEKLRLMPYQEQDLLREHYQKKFKNSTYSRSSVDVLYTFPANC 180
QY 181 GDLIFGLNALIRTDALQWSSNAQLLDYCSSKGYNISWELGNEBNSFLKKADIFINGS 240
DB 181 GDLIFGLNALIRTDALQWSSNAQLLDYCSSKGYNISWELGNEBNSFLKKADIFINGS 240
QY 241 QGGEPIQHLKLRSTFNKALYGPVQOPRRKTAKMKSFLKAGGEVIDSTVTHHYYL 300
DB 241 QGGEPIQHLKLRSTFNKALYGPVQOPRRKTAKMKSFLKAGGEVIDSTVTHHYYL 300
QY 301 NGRTATREDPLNPVDLDFISSVQKVFQVVESTPGKTYMLGETSSAYGGAPLISDTFA 360
DB 301 NGRTATREDPLNPVDLDFISSVQKVFQVVESTPGKTYMLGETSSAYGGAPLISDTFA 360
QY 361 AGFMWLDKLGISARMGIEVVMQVFFGAGNYHLVDENFDPLPDYWLSTLLFKKLVTGTYLM 420
DB 361 AGFMWLDKLGISARMGIEVVMQVFFGAGNYHLVDENFDPLPDYWLSTLLFKKLVTGTYLM 420
QY 421 ASVQSSKRRKRLRVYLHCTNDNPRYKGGDLTYAINLHVNTKYLRPPFSNKQVDKYL 480
DB 421 ASVQSSKRRKRLRVYLHCTNDNPRYKGGDLTYAINLHVNTKYLRPPFSNKQVDKYL 480
QY 481 RPLGPHGLSKSVQNLGLTLKKVDDQTLPLMEKPLRPGSSISGLPAFSYSPFVIRNAKYA 540
DB 481 RPLGPHGLSKSVQNLGLTLKKVDDQTLPLMEKPLRPGSSISGLPAFSYSPFVIRNAKYA 540
QY 541 ACT 543
DB 541 ACT 543

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2005, 11:09:57 ; Search time 113.5 Seconds
(without alignments)
1850.316 Million cell updates/sec

Title: SEQ2B
Perfect score: 2841
Sequence: 1 MLIRSKPALPPIMLLIGP.....LPAFSYSPFVIRNAKVAACI 543

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues
Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2841	100.0	543	2	AAV17082 Human hep
2	2841	100.0	543	4	AAB86206 Human hep
3	2841	100.0	543	7	ADD18950 Human dis
4	2841	100.0	543	8	AdK52086 Human atc
5	2841	100.0	543	8	AdM48759 Human hpa
6	2841	100.0	543	8	AdN05074 Antipsoi
7	2841	100.0	543	8	AdN04902 Antipsoi
8	2841	100.0	543	8	AdQ80372 Heparinas
9	2841	100.0	543	8	AdR88210 Human pre
10	2841	100.0	543	8	AdP25079 PRO poly
11	2841	100.0	588	2	AAV30124 A human p
12	2838	99.9	543	2	AAV02345 A human h
13	2838	99.9	543	3	AAV57590 Human hep
14	2838	99.9	543	3	AAB08849 Amino aci
15	2838	99.9	543	3	AAV52990 Human hep
16	2838	99.9	543	3	AAV57635 Human hep
17	2838	99.9	543	5	ABO7813 Human hep
18	2838	99.9	543	7	ADG88800 Human hpa
19	2838	99.9	543	8	ADL16379 Human hep
20	2838	99.9	543	8	AdM48716 Human hpa
21	2838	99.9	592	2	AAV02346 A human h
22	2838	99.9	592	3	AAB08850 Amino aci
23	2838	99.9	592	7	ADG88804 Human SK-
24	2838	99.9	592	8	ADL16383 Human SK-
25	2838	99.9	592	8	AdM48720 Human SK-

26	2835	99.8	543	8	AD063831 Human hep
27	2835	99.8	543	8	AD063823 Human hep
28	2835	99.8	543	8	AD063832 Human hep
29	2835	99.8	543	8	AD063822 Human hep
30	2829	99.6	543	4	AAB88361 Human mem
31	2829	99.6	543	8	AD063824 Human hep
32	2820	99.3	545	6	ABP56822 Human hep
33	2820	99.3	545	7	ADL16012 Human G-c
34	2820	99.3	545	8	ADL193951 Human G-c
35	2767	97.4	530	2	AAV34173 Human pre
36	2740	96.4	532	2	AAV17083 Seg ID No
37	2676.5	94.2	527	5	ABO7815 Chicken b
38	2676.5	94.2	527	7	ABW02018 ChimERIC
39	2670.5	94.0	527	8	AD063825 ChimERIC
40	2670.5	94.0	527	8	AD063826 ChimERIC
41	2664.5	93.8	527	8	AD063827 ChimERIC
42	2149	75.6	535	3	AAB08851 A murine
43	2149	75.6	535	5	ABO7811 Mouse hep
44	2149	75.6	535	7	ADG88834 Mouse hpa
45	2149	75.6	535	8	ADL16413 Mouse hep

ALIGNMENTS

AAV17082	AAV17082 standard; protein; 543 AA.
AC	AAV17082:
DT	21-JUL-1999 (first entry)
XX	Human heparanase enzyme.
DE	Heparanase; endoglucuronidase; heparan sulfate proteoglycan; enzyme; metacastis; angiogenesis; wound healing; angioplasty-induced restenosis; KX
XX	arteriosclerosis; atherosclerosis; inflammation; tissue development; KX
XX	human; HSPG. KX
OS	Homo sapiens. KX
XX	WO921975-A1. KX
PN	06-MAY-1999. KX
PD	28-OCT-1998; 98WO-AU000898. KX
PF	28-OCT-1997; 97AU-0000062. KX
XX	09-DEC-1997; 97AU-00000812. KX
PR	(AUSU) UNIV AUSTRALIAN NAT. KX
PA	Freeman CG, Hulett MD, Parish CR, Hamdorf BJ, KX
PI	WPI, 1999-312956/26. KX
XX	N-PSDB; AAX37259. KX
DR	Polynucleotides encoding mammalian endoglucuronidases, especially KX
XX	heparanases, useful to promote wound healing. KX
PT	Claim 6; Page 69-73; 112pp; English. KX
PS	The invention relates to nucleic acid sequences that encode heparanase KX
XX	enzymes having endoglucuronidase activity. Recombinant heparanases are KX
XX	capable of removing the HS side chain from heparan sulfate proteoglycan KX
CC	(HSPG). Sulfated oligosaccharides, sulphonates or HSPG can be used to KX
CC	inhibit heparanase, this is useful for treatment of a physiological or KX
CC	medical condition associated with elevated heparanase activity, such as KX
CC	metastasis, angiogenesis, wound healing, angioplasty-induced restenosis, KX
CC	arteriosclerosis, atherosclerosis and inflammation. The human, murine and KX
CC	rat heparanases can be used to enhance wound healing, especially KX
CC	associated with tissue development and repair. The conditions mentioned KX

This Page Blank (uspto)

RESULT 12

S00652

phosphoribosylamine-glycine ligase (EC 6.3.4.13) - fission yeast (*Schizosaccharomyces pombe*)
 NAlternate names: AIRase, aminimidazole ribotide synthetase, GARSae, glycylamide rib N-Transferase
 NContains: phosphoribosylamine-glycine ligase (EC 6.3.4.13); phosphoribosylformylglycin C;Species: *Schizosaccharomyces pombe*
 C>Date: 07-Sep-1990 #sequence revision 28-Oct-1994 #text_change 09-Jul-2004
 C/Accession: S00652; T40496; T40422

R/McKenzie, R.; Schuchert, P.; Kilday, B.
 Curr. Genet. 12, 591-597, 1987

A/Title: Sequence of the bifunctional ade1 gene in the purine biosynthetic pathway of *Ch*

A/Reference number: S00652; M01D:89003164; PMID:3502942

A/Accession: S00652

A/Molecule type: DNA

A/Residues: 1-788 <MCK>

A/Cross-references: UNIPROT:P20772; EMBL:X06601; NID:g4903; P1DN:CAA29820.1; PID:g4904
 R/Mood, V.; Rajandream, M.A.; Barrell, B.G.; Lauber, J.; Hiltbert, H.; Duesterhoeft, A.
 submitted to the EMBL Data Library, February 1998

A/Reference number: Z21910

A/Accession: T40496

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-788 <MOO>

A/Cross-references: EMBL:AL021730; P1DN:CAA16823.1; GSPDB:GN00067; SPDB:SPBC4C3.02c
 A/Experimental source: strain 972h; cosmid c4C3

R/Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, March 1999

A/Reference number: Z21928

A/Accession: T40422

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 604-788 <SEE>

A/Cross-references: EMBL:AL035655; P1DN:CA38660.1; GSPDB:GN00067; SPDB:SPBC405.01

A/Experimental source: strain 972h; cosmid c405

C/Genetics:

A/Gene: ADE1; SPDB:SPBC405.01

A/Map position: 2
 C/Superfamily: Saccharomycetes cerevisiae AD5 multifunctional protein; phosphoribosylamir

F/5-4/5/Domain: phosphoribosylamine-glycine ligase biosynthesis <PGL>

F/439-767/Domain: phosphoribosylformylglycinamide cyclo-ligase homology <PFL>

Query Match 3.7%; Score 104.5; DB 1; Length 788;
 Best Local Similarity 27.7%; Pred. No. 14;

Matches 70; Conservative 36; Mismatches 114; Indels 33; Gaps 11;

297 HYLLNGRTATRE--DFLNPV--LDIFISSVQKVFQVEST--RPGKVMLGESSAY---- 348

DB 424 HHALNPKRKTRILTYENGVSVDNNGNEFVQRIKDLVSTRRGADADIGGFGIFDLKQ 483

QY 349 -GGGAPLL--SDPPAAGFMMLDKGLSAR--MGIEVVMKQVFPAGNYHLVDENFDL--P 402

DB 484 AGNNDLLVASVDGVSKLLIALSLNKHDTVGIDLVAMNV-----NDLVVQGAELIFL 537

QY 403 DVMLSLFLKVLGTGYKVLMAVQSGSKRKRLRVYLHCTNTDPRYKGGDLTLVAILNANTR 462

DB 538 DIFATISDLKSTSTVEGVKCKQACALVGGESSEMPGLYHDHYDANGSVGAVSR 597

QY 463 YRLPLVPFNSNKQVDTKYLRLPGHGLLSKSVQNLGLTL--KMYD-----DQTLPLMEKPL 516

DB 598 DDLPRPESFSSKDDILL-----GLASDGVHNSNGSLVAKIYEVSDLEYTSVCPMDKAV 650

QY 517 RGGSSGLAPFSY 529

DB 651 RUGDSLILPFRY 663

RESULT 13

F70411

adenylosuccinate synthetase - *Aquifex aeolicus*

C/Species: *Aquifex aeolicus*

C/Date: 08-May-1998 #sequence revision 08-May-1998 #text_change 09-Jul-2004

C/Accession: F70411

R/Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Oye

V. Nature 392, 353-358, 1998

A/Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.

A/Reference number: A70300; M01D:98196666; PMID:9537320

A/Accession: F70411

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-432 <AQF>

A/Cross-references: UNIPROT:O67321; GB:AE000733; NID:g2983720; P1DN:AA07286.1; PID:g2983

A/Experimental source: strain VFS

C/Genetics:

A/Gene: purA

C/Superfamily: adenylosuccinate synthase

Query Match 3.7%; Score 104; DB 2; Length 432;
 Best Local Similarity 23.9%; Pred. No. 6.2;

Matches 96; Conservative 39; Mismatches 128; Indels 138; Gaps 22;

15 LLLIGPLSPGALPRPAQDVVDLD-----FTQBPMLVSPS 55

DB 51 LHLPLTGLHGHVKVIAQGM--VDLEVLHKEVKLEKGIYKRLTSDRAHLWMPY 109

QY 56 FLVITDANIATDPRLILGSPK--LRTLAGLSPAY--REGTKTDFLIDPKKESTF 112

DB 110 H-----KLDLSEFKKKIGITLAKIGAVYFKG--RKGRISDLKDEKRF 154

QY 113 EERSTWQSOVNOIDCKYSGISPPDVEEK-----LRLEMPYQOLLRHRYQKKFNSTY 165

DB 155 ----YTLLEDNDLDFV-----NICEKVCERFDLDINOIYEOI-----RYEEEFKEN- 199

QY 166 SRSVDVLTFPANCSGLDIFGLNLLRTADL-----QWNSNAQLLDYCSSKGVNISM 221

DB 200 ---VDLRFPTQKGSVLFEGAGTLDLDVMTGTPYVTSNNSAL-----GLSNG 246

QY 222 LGNBNPSFLKADIFING-----SQI--GDFIQLHKLKSTFKNAKLYG 265

DB 247 TGMPPRYF--SDAPFLGAVAKAYTRVGGPPPTBELKGEGERLRL-----CGERG 295

QY 266 PDVGQPRR--KTAKMLKFLAAGEVIDSVTHHYLLNGRTATREDFLNP----- 313

DB 296 STYGRPRRCGMIDLVALKTAQVONG-----LDGFVITKLDVDFDEVKCVVA 343

QY 314 ----DVLDIFISSVQKVFQV--VESTPRGKVMLGESSA 347

DB 344 YELDGEVIDYFPASYSSELIRKVPVYKTLKG---WKKSTXGA 381

RESULT 14

D97065

transketolase [imported] - *Clostridium acetobutylicum*

C/Species: *Clostridium acetobutylicum*

C/Date: 14-Sep-2001 #sequence revision 14-Sep-2001 #text_change 09-Jul-2004

R/Accession: D97065

R/Molling, J.; Breton, G.; Omechenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

J. Bacteriol. 183, 4823-4838, 2001

A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clo*

A/Reference number: A96900; M01D:21359325; PMID:21359325

A/Accession: D97065

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-796 <KUR>

A/Cross-references: UNIPROT:Q97JF3; GB:AE001437; P1DN:AAK7931.1; PID:g15024275; GSPDB:GT

A/Experimental source: *Clostridium acetobutylicum* ATCC824

C/Genetics:

A/Gene: CAC1343

C/Superfamily: phosphoketolase

Query Match 3.7%; Score 104; DB 2; Length 796;
 Best Local Similarity 21.6%; Pred. No. 15;

Matches 77; Conservative 51; Mismatches 114; Indels 114; Gaps 18;

```

QY          416 TKVLMAV 423
      | | :
Db          401 LKSLHQT 408

RESULT 11
E91031
Probable outer membrane protein ECs3221 [imported] - Escherichia coli (strain O157:H7, su
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #ext_change 09-Jul-2004
C:Accession: E91031
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gsawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
A:Reference number: A99625; MUID:21156231; PMID:11258796
A:Accession: E91031
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-879 <HAV>
A:Cross-references: UNIPROT:Q8XCP4; GB:BA000007; PIDD:BA936644.1; PIDD:913362691; GSPDB:GN
A:Experimental source: strain O157:H7, Substrain RIMD 0509952
C:Genetics:
A:Gene: ECs3221

Query Match          3.7%; Score 105.5; DB 2; Length 879;
Best Local Similarity 19.9%; Pred. No. 14; 208; 241; Gaps 33;

```

[illegible][illegible]

QY 256 -----STFKN--AKLYGPDVGOPRRKTAFLKAGSEV-----IDSV 293
 DB 341 DYKDFHIGVITKYLNSGFKNNYSTLMTSSKTQDRKSHNMSSILDDGNKIGMHPISIDE- 399
 QY 294 TWHHYYLNG-----RTATREDFLNPVDLIDITSSVQKVPQVSESTR--PGKK 338
 DB 400 -YSHFIDNDEPLMRDKVPKPYTNEGTPTPDASAIFFDS--HKIYAIISILRYLPEKR 454

RESULT 7

hypothetical protein M0670 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004

C:Accession: F64383
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
 A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
 A:Reference number: A64300; MUID:96337999; PMID:8688087

A:Accession: F64383

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-356 <BUL>

A:Cross-references: UNIPROT:Q58084; GB:U67514; GB:L77117; NID:92826304; PIDN:AB98664.1;
 C:Genetics:

A:Map position: REV596956-595886

A:Start codon: GTG

Query Match 3.9%; Score 110.5; DB 2; Length 356;
 Best Local Similarity 21.0%; Pred. No. 1.5; Indels 105; Gaps 15;

Matches 82; Conservative 49; Mismatches 155; Indels 105; Gaps 15;

QY 126 ICKY-----GSIPDVEKRLLEMPYQEQLLREHYQKKFKNSTYSSVD----- 171
 DB 14 IRRKMIYGNKBEKIDKERLKE-----LKEHVLVETEDGTYYLKADEBEMHSHV 66

QY 172 -----VLYTFANCGLDLIFGLNALRLTADLQMNSSNAQLLDYCSKGYNISMEIGNEP 226
 DB 67 GALKKAIYFAPKS-----KIDL-----SNPR-VLDLCSGMVMAIAALHYNK 109

QY 227 NSFLLKADIF-----INGSQLGDFIQLHKLKSTFKNAKLYGPDV 268
 DB 110 NAEIDMVEICEVLPITLFDLPYKHEIKDKVRYFLANKIGIEKSDYDINLY--V 166

QY 269 GQPRRTAFLKSLKAGGEVIDSVTHHHYLLNGRTAT--REDFLNPVDLIDITSSVQK 326
 DB 167 GDAKRFIIKSDKY-----NVVEHDAFSPKRDPTLYTDFL-----KEI 205

QY 327 PQVVESTRRGKVMGLGETSAYGGAPLLSDTFAAFMMDLGLSARMGIEVMQVFP 386
 DB 206 YKRMEN--GVLI-----SYSSALPFRSALVDCFFVISEKSVKRGKITLAYNPNF 256

QY 387 GAGNYHVLVDENFD-----PLPDYWLILFKLVGTGYKLAMSVQSKRKLRYVLACIN 439
 DB 257 KENRINEVERVIALSVIALPYRDELTLSTKDKIIDREBRKLEKLIKIKYUJSTQ 316

QY 440 TQNPYKEDDLTYA--INLHNTYKYLRLPY 468
 DB 317 IKKGNPEELIKQEDLNSSEIKKGRKF 347

RESULT 8

beta-fructofuranosidase (EC 3.2.1.26) - fava bean

C:Species: Vicia faba (fava bean)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

C:Accession: T12094

R:Weber, H.; Borisjuk, L.; Heim, U.; Buchner, P.; Wobus, U.

Plant Cell 7, 1835-1846, 1995

A:Title: Seed coat-associated invertases of fava bean control both unloading and storage

A:Reference number: Z17416; MUID:96093423; PMID:8535137

A:Accession: T12094

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-575 <WEB>

A:Cross-references: UNIPROT:Q43855; EMBL:Z35162; NID:9861154; PIDN:CAA84526.1; PID:986115

A:Experimental source: cv. Fibro, seed coat

A:Genetics:

A:Gene: CMINV1

C:Superfamily: beta-fructofuranosidase

C:Keywords: cell wall; glycoprotein; glycosidase; hydrolase

Query Match 3.8%; Score 107.5; DB 2; Length 575;
 Best Local Similarity 21.1%; Pred. No. 5.2;

Matches 71; Conservative 49; Mismatches 107; Indels 109; Gaps 19;

QY 46 QEPHLIVS-----PSFLSVITDANLATDPFLLILSGPKRTLARGIS-----P 89
 DB 228 KPHISAKRTGMECPDFPVSLEKNGID--LSMMGNVHVLKNSLIDITREYTYTG 285

QY 90 AYLR-----FGTKTDP-----LIFDPKKESTFEERSYV-----QSQVNG 124
 DB 286 TYLQNDKTIIPDKTSEDGKGLRDIYGNFYASKSFDPYK--NRIITGMANESDTKE 341

QY 125 DICKYG-----SIPDV-----EKKLRLEMPYQEQLLR-----EHYQKKFKNSTYSSSV 170
 DB 342 DQVKGGMAGIQAIPTVWMLDSSRRQLR-QWVPELNLRLGKQVEMKNRLKXGGY----L 396

QY 171 DVLTFPANCGLDLIFGLNALRLTADLQMNSSNAQLLDYCSKGYNISMEIGNEPFTL 230
 DB 397 EVKGITASQADVEVTFSSLDKAEAFDPNMENAE--DLCKQKSKVGVG--PFGLL 451

QY 231 KKADIFINSQLEPDIQ-----HKL-----RKSTFKNAKLYGPDV-----DV 268
 DB 452 TLA-----SKLEETSVYFRRVFKAAKHLKMSDASSLSNRLYPSFAGFVNVDL 505

QY 269 GQPRRTAFLKSLKAGGEVIDSVTHHHYLLNGRT 304
 DB 506 GNNKLSLRSI-----IDHVSVEFGVGKT 531

RESULT 9

microtubule-associated protein sm4 - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: T38446; T00012

R:McDougall, R.; Wood, V.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, December 1997

A:Reference number: Z21793

A:Accession: T38446

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-670 <MCD>

A:Cross-references: UNIPROT:Q42667; EMBL:AL009227; PIDN:CAA15832.1; GSPDB:GN00066; SPDB:6

A:Experimental source: strain 972h; cosmid c27D7

R:Yamashita, A.; Watanabe, Y.; Yamamoto, M.

Genes to Cells 2, 155-166, 1997

A:Title: Microtubule-associated coiled-coil protein Sm4 is involved in the meiotic devel

A:Accession: T00012

A:Reference number: Z14042; MUID:97311255; PMID:9167972

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-670 <YAM>

A:Cross-references: EMBL:AB000269; NID:93341860; PIDN:BAJ31857.1; PID:93341861

A:Genetics:

A:Gene: sm4; SPAC27D7.13c

A:Map position: 1

Query Match 3.7%; Score 106; DB 2; Length 670;
 Best Local Similarity 21.4%; Pred. No. 8.5; Indels 86; Gaps 15;

Matches 80; Conservative 58; Mismatches 149; Indels 86; Gaps 15;


```

QY 411 KLVGTGYLMASVQSKRR-----KLRYLHCTNTDNPRYEGDGLTYAINLHNT 461
DB 337 KRLIPKVLAVAVAGIQKRPGRVIRDLKRIYAHCTNNHNNYRGSTLTFLINHR 396
QY 462 KYLRLLPYPPSKNOVDKYLRPLGPHGLLSKVQNLGLTKMVDQTLPLMEKPLRPGSS 521
DB 397 KIKIKLAGTLRDKLIVHQLYLLQPYGEGELKSKSVQLNGQPLVMVDDGLTLPKRPRLAGRT 456
QY 522 LGLPAFSYFFYIRNAKVAAC 542
DB 457 LVIPPTMGFFVKKVNNALAC 477

RESULT 2
T45608
Hypothetical protein F13G24.30 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C/Accession: T45608
R/Bevan, M.; Van Der Schueren, J.; Chung, Y.J.; Voel, M.; Robben, J.; Volckaert, G.; Be
submitted to the Protein Sequence Database, December 1999
A/Reference number: Z23009
A/Accession: T45608
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-521 <BEV>
A/Cross-references: UNIPROT:Q9SDA1; EMBL:AL133421
A/Experimental source: cultivar Columbia; BAC clone F13G24
C/Genetics:
A/Map position: 5
A/Introns: 53/3; 66/1; 127/2; 177/1; 256/1; 319/2; 361/2; 394/3
A/Note: F13G24.30

Query Match 14.7%; Score 417; DB 2; Length 521;
Best Local Similarity 29.2%; Pred. No. 4.3e-23;
Matches 154; Conservative 68; Mismatches 184; Indels 122; Gaps 24;

QY 75 LSPKRLTLAGSPAYLRFPGTKTDFLFDPKKSTPERSWQGVQVNDICVKYSIRP 134
DB 55 LTRPLTKAIKAFKPLRIRIGSLQDVQYDVGNLKT-----PCR----- 94
QY 135 DVEBKRLRWPYOEQLLREHYOKKFKNS--TYSRSV-----DVLTYFANCSGLDIF 186
DB 95 -----PFQKM-----NGLRGFSKGCGLMKRMWDELFLATATGAVTF 132
QY 187 GLNALIRTDQ-----WSSNAQLLDYSSKGYNT-SWELGNEPNSFLKADIFIN 238
DB 133 GLNALRGHKLKRGKAMGAMDHINTQDFLNTYVSKGYVIDSWFEGNELSG--SGVGASYS 190
QY 239 GSQLGDFIQHLKLRKSTFKNAKLYGPRVGP-----RRTKAKMLKSLFKAGGEVIDSV 293
DB 191 AELVGDLLVLKDVINK-VYKSWLHKPLVAPGGGYEQOWYTKLEI---SGPSVADV 246
QY 294 TWHHYVINGRT--ATREDFLNPVDLIDIFISSVQKV---QVVESTRPGKWLGETSSA 347
DB 247 THHIVYLSGSDNPALYKTKMDPS-----YLSQVSKTKVDNQTQEGHPMSPVGSSEA 302
QY 348 YGGGADLSDTPAAGFMWLDKGLSARWGI EVVMRQVFFGAGNYHLVDE-NEDPLDYWL 406
DB 303 YNSGGRHVSDTFIDSFMYLDQLGMSARHNTKYVCRQLVG-GPYGLLEKGTFFVNDYYS 361
QY 407 SLTFKLVGTGYLMASVQSKRRKLRVYLHCTNTDNPRYEGDGLTYAINLHNTKYL-- 464
DB 362 ALLMRLKMGVLA VOTDGP--QLRVYAHCSK-----GRAVTLTLNLNSQSDFTVS 413
QY 465 -----RLPYFS--NKQVDKYLRLP---LGPHG--LLSKSVOL 495
DB 414 VNSGINVVLNABSRKKKSLDLTLKRFPSWIGSKASGYLNRBEYHLLTPENGVLRSKTMVL 473
QY 496 NGTLTKMVDQTLPLMEKPLRP-GSSLGIPAFSYFFYIRNAKVAAC 542
DB 474 NKSILKPTATGDIPLSLEPVLRSVNSPLNVLPLSMGFIVLPNPDASAC 520

```

```

RESULT 3
T01953
Hypothetical protein T2L5.6 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 09-Jul-2004
C/Accession: T01953
R/Giesel, C.; Smith, A.; Le, T.
submitted to the EMBL Data Library, October 1998
A/Description: The sequence of A. thaliana T2L5.
A/Reference number: Z14470
A/Accession: T01953
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-190 <GEI>
A/Cross-references: UNIPROT:O82604; EMBL:AF096371; NID:93695386; PID:93695392
A/Experimental source: cultivar Columbia
C/Genetics:
A/Map position: 4
A/Introns: 36/2; 69/3
A/Note: T2L5.6
C/Superfamily: Arabidopsis thaliana hypothetical protein T2L5.6

Query Match 6.0%; Score 169.5; DB 2; Length 190;
Best Local Similarity 27.8%; Pred. No. 2.4e-05;
Matches 54; Conservative 34; Mismatches 57; Indels 49; Gaps 9;

QY 382 ROVPPGAGNYHLVD-ENFDPLPDYWLSLFKKLVGTGYLMASVQSKRRKLRVYLHCTVT 440
DB 12 RQSLIG-GNYGLLNTFTFPNDYYSALIMWOLMRKALFTTFSGTK--KIRSYTHCA-- 66
QY 441 DNPYKRGDGLTYAINLHNV-----TKYLRLLPYPPSKNOVDKYLRPL 483
DB 67 ---RQSKG-ITVLLMLNDVTTTVAKVLLNNSFSLRHKKMK-----SYKRASSQLFG-- 115
QY 484 GPHGLL-----SKSVQNLGLTLKMWDDQTLPLMEKPLRPGSSLGIPAFS 528
DB 116 GNGVIGQEEHYHLTKADGSLHSQTMNLGNLQVNSMGDLPRIEPLHINSTEPIITAFYS 175
QY 529 YSFVIRNAKVAAC 542
DB 176 IVFVIMRNVVPAC 189

RESULT 4
T49648
Hypothetical protein B8B20.20 [imported] - Neurospora crassa
C/Species: Neurospora crassa
C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C/Accession: T49648
R/Schilte, U.; Aign, V.; Hohnsiegel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A/Reference number: Z25022
A/Accession: T49648
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-2298 <SCH>
A/Cross-references: UNIPROT:Q96U00; EMBL:AL355933; GSPDB:GNO0116; NCSP:B8B20.20
A/Experimental source: BAC clone B8B20; strain OR74A
C/Genetics:
A/Gene: NCSP:B8B20.20
A/Map position: 6
A/Introns: 426/3

Query Match 3.9%; Score 112; DB 2; Length 2298;
Best Local Similarity 19.3%; Pred. No. 19;
Matches 114; Conservative 79; Mismatches 190; Indels 208; Gaps 28;

QY 78 PKRLTLAGLSPA-----YLRFGGTYKDFLFDPKKSTPERSWQGVQV-NQDIC 127
DB 1447 PVDLIERLTIPSNHAKACLINIRAMNQLARLVVNSGSGASFRPFTIRNNVFNQIID 1506

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2005, 11:19:33 ; Search time 27 Seconds
(without alignments)
1935.026 Million cell updates/sec

Title: SEQ2B
Perfect score: 2841
Sequence: 1 MLRSKPLPPPLMLLLGP.....LPAPSYSPFVIRNAKVAACI 543

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir79:*
2: pir1:*
3: pir2:*
4: pir3:*
5: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	893.5	31.5	480	2 JC7506	heparanase protein
2	417	14.7	521	2 T45608	hypothetical prote
3	169.5	6.0	180	2 T01953	hypothetical prote
4	112	3.9	2238	2 T49648	hypothetical prote
5	111	3.9	670	2 T10666	hypothetical prote
6	111	3.9	688	2 S32961	hypothetical prote
7	110.5	3.9	356	2 F64383	hypothetical prote
8	107.5	3.8	575	2 T12094	beta-fructofuranos
9	106	3.7	670	2 T38446	microtubule-associ
10	105.5	3.7	411	2 S74760	hypothetical prote
11	105.5	3.7	879	2 E91031	probable outer mem
12	104.5	3.7	788	1 S00652	phosphoribosylamin
13	104	3.7	432	2 F70411	adenylosuccinate s
14	104	3.7	796	2 D97065	transketolase (imp
15	104	3.7	2013	2 A11489	probable peptidogl
16	103.5	3.6	500	2 D87541	tetracycline (1
17	103.5	3.6	676	2 AF1153	transcription anti
18	102	3.6	879	2 F85875	probable fibrinoly
19	102	3.6	897	2 G02529	dynein heavy chain
20	102	3.6	4644	1 A38905	dynein heavy chain
21	101.5	3.6	746	2 T46821	siderophore recept
22	101.5	3.6	746	2 A95420	Rhizobactin r
23	101	3.6	594	2 A82913	hypothetical prote
24	100.5	3.5	604	2 E75119	hypothetical prote
25	100.5	3.5	687	2 F85188	retrotransposon 11
26	100.5	3.5	847	2 AG1001	nitrite reductase
27	100.5	3.5	1314	2 S19488	probable membrane
28	100.5	3.5	1734	2 A41101	phorbol ester-bind
29	100	3.5	578	2 B89045	protein B0238.7 [1

30	100	3.5	654	2 T14202	NADH2 dehydrogenas
31	99.5	3.5	587	2 S36231	beta-fructofuranos
32	99.5	3.5	989	2 AE2140	toxin secretion AB
33	98.5	3.5	629	2 C64180	hypothetical prote
34	98.5	3.5	804	2 G71546	probable DNA gyrase
35	98	3.4	465	2 T19113	hypothetical prote
36	98	3.4	644	2 A97268	methionyl-tRNA syn
37	98	3.4	716	1 C60008	RNA-directed RNA p
38	97.5	3.4	511	2 S61166	probable membrane
39	97	3.4	379	2 A69974	cystathionine gamm
40	97	3.4	437	1 A48061	translational releas
41	97	3.4	726	2 C66085	catalase, hydrotrop
42	97	3.4	726	2 G91237	hydrotropoxidase HP
43	97	3.4	760	2 T34414	hypothetical prote
44	96.5	3.4	510	2 H69893	conserved hypothet
45	96.5	3.4	621	2 A95250	choleine binding pr

ALIGNMENTS

RESULT 1

JC7506
heparanase protein 2a - human
C/Species: Homo sapiens (man)
C/Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #ext_change 09-Jul-2004
C/Accession: JC7506
R/McKenzie, E.; Tyson, K.; Stamps, A.; Smith, P.; Turner, P.; Barry, R.; Hircock, M.; Pat
Biochem. Biophys. Res. Commun. 276, 1170-1177, 2000
A/Title: Cloning and expression profiling of Hpa2, a novel mammalian heparanase family me
A/Reference number: JC7506
A/Accession: JC7506
A/Molecule type: mRNA
A/Residues: 1-480 (MWCK)
A/Cross-references: UNIPROT:Q9HB39; GB:AF282885
C/Comment: This protein, a intracellular membrane-bound enzyme, has biological and therai
Chetraples.
C/Genetics:
A:Gene: hpa2a
A:Map position: 10q23-10q24
C/Keywords: heparin binding; membrane bound

Query Match	31.5%	Score 893.5	DB 2	Length 480
Best Local Similarity	35.8%	Pred. No. 1,4e-58		
Matches	201	Conservative	75	Mismatches 146; Indels 139; Gaps 9
QY	20	PLGSLSPGAL-----PRPA-----OAQDVDDPFTQEPHLVSPS	55	
DB	18	PPACLAAPGALYALALLHLSSQAGDRRLPVDRRAAGLKEKTLILLDVSTKAPRVNEN	77	
QY	56	FLSVTIDANLADPRLILIGSPKLTTLARGLSPAYLRFQGTWDFLIF---DPKKEST	111	
DB	78	FLSLQDPSLIHD-GWLDFLSSRKLTVLARGLSPALFRGKRTDPLQONLNNPAKSR-	135	
QY	112	FEERSYWGSCVONODICKYSSIPPDVEKRLRLEWYQELLRHHYQKKRKNSTYSSSVD	171	
DB	136	-----GGPGPD-----YYLKNYE-----	148	
QY	172	VLTAFNCGSLDLIFGLNLLRLADQLQWNSNAQLLDYCSSKGYNISWELGNEPNSFLK	231	
DB	149	-----DEPNMYRT	156	
QY	232	KADIFINGSQLGDFIQLHKLRLK-STFKNAKLYGPDVQPRRKTKAMLSFLKAGGEVI	290	
DB	157	MGRRAVNSQLGKDIYQLKSLIQPIRITYSRASLYGNIRPRKNVIALDGFMKVAGSTV	216	
QY	291	DSYTMHYTLNGTARREDPLNDVDIFLSSVQKXFOYVESIRPKKXWLGSTSSAYCG	350	
DB	217	DAVTWQHCYIDGRVAVKVMDFLKTRLDLTSDQIRKIQKVVNTYTPKKKTLLEGVVTSSAG	276	
QY	351	GAPLSDTFAAGFMWLDKLSARMGIEVVMROVFGAGNYHLVDNFPDLPDYMVSLIF	410	
DB	277	GTMNLSDSYAGFGLMINTLGLMLANQGLIDVYIRHSFFDHGHNHLVDQNFPLPDYMSLILY	336	

This Page Blank (uspto)

```

QY 106 PKKESTFEERSYQWQONODICKYGISPPDVEBKLRLMPYQEBLLRHHYOKKFKNSY 165
DB 100 -SKENI-----PISCHNCSYKSYPRSLCO--LIERPC-----NKKKFLPFIIM 140
QY 166 SRSVDVLTYPANCSGLDILFGNALILFTADLQWNSNAQLLDYCSSKGYNISWELGNE 225
DB 141 TGNEMNOINDFCRKTMLKLLFSLNMLRP-NHGMENKARELIEFKHKQYALMDQLGNE 199
QY 226 PMSFLKADIFINGSLGSDFIQLHKLLRKSFTFKNAKLYGPDVGOP--RRKTAKMLKSF 282
DB 200 PMSFGVFNESVTPQILAKDFEKLRLNHNNGVRHSLIAGPDTTRPQHRPRLCKYMIFF 259
QY 283 LAAAGVIVSVTHNYLNGRTATREDFLNPDVLDIFISSVOKPQVVESTPGKKV--M 340
DB 260 LGNGSHYIVRSMHQYLLNSKTAKLEDPMWPEFEDL--ROQIETMQOTKKYKIIPIW 316
QY 341 LGETSAYGGGAPLLSDTFPAAGFMWLDKGLSARMGIEVVMROVFFGAGNYHLVDENPD 400
DB 317 LSETSSYGGGAPGLSNTYAGSPMLDKGLSANKNISYIRQSFIG-GYSLVDENLKP 375
QY 401 LPDYWLSLFKKLVGTRVMASVQSSKRRKRLVYLHCTNTDNPYKE--GDULTYAIN-- 456
DB 376 LPEMWISVLYKKLVGKVL--QVQCNCSPFORLYIHCTNR--KYNDTSAYVLGYVNL 430
QY 457 -----LHN-----VTXYLRLPYPSNKOVDKYLLRPLGPHGLSKSVQLNGLT 500
DB 431 MAYARFFLANGTALHGDLLIHERY-ISAESNNK-----SKITLLANGWFL 474
QY 501 KMWDDQTLPELMEKPLRPGSSSLGPAFSSYFFVIRAKYAAAC 542
DB 475 YV--ESNLHLRPNIRHYGRYVSLPPYSIGFWVIKKTSTIVC 514

```

RESULT 15

```

ID Q9SDA1 PRELIMINARY; PRT; 521 AA.
AC Q9SDA1;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Hypothetical protein F13G24.30.
GN Name=F13G24.30;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; core eudicots; rosids;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eudicots II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Van Der Schueren J., Chuang Y.J., Voet M., Robben J.,
RA Volckaert G., Bancroft I., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DB EMBL, AL133421, CAB62595.1; -
DR PIR; T45608; T45608.
DR InterPro; IPR005199; Glyco_hydro_79N.
DR InterPro; IPR001254; Peptidase_S1.
DR Pfam; PF03662; Glyco_hydro_79n; 1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 521 AA; 57831 MW; 07D8664AAB305CC2 CRC64;

```

Query Match 14.7%; Score 417; DB 2; Length 521;
 Best Local Similarity 29.2%; Pred. No. 7.4e-23;
 Matches 154; Conservative 68; Mismatches 184; Indels 122; Gaps 24;

QY 75 LGSPEKRLTLAGLSPAYLRFGCTKDTPLFPKKESTFEERSYQWQONODICKYGSIPP 134
 DB 55 LTRPLLTVAIKAFKPLRIRIGSLQDVIYIVGNLKT-----PCR----- 94

```

QY 135 DVEEKLRLMPYQEBLLRHHYOKKFKNS---TYSRSSV-----DVLTPANCSGLDILF 186
DB 95 -----FPQCK-----NSGLFGSKGCLHMKRNDLNSFLATGAVTF 132
QY 187 GINALLRTADLQ-----MNSNAQLLDYCSSKGYNI--SWELGENENSLFKADIFIN 238
DB 133 GINALRGHKLKGKAMGAMDHINTQDFLNTYVSKGYVIDSMWFENELSG--SGVGASVS 190
QY 229 GSQAGEDFIQLHKLLRKSFTFKNAKLYGPDVGOP-----RRKTAKMLKSFKAGGEVIDSV 293
DB 191 AELYGKDLVLKDVINK-VTKNSWLHKFLLVAFGAFYEQQWYTKLLEI---SGPSVDVY 246
QY 294 TWHHYLLNGRT--ATREDFLNPDVLDIFISSVQVF---QVESTPGKKYWLGETSA 347
DB 247 THHITNLGSGNDPALVKKIMDS-----YLSQVSKTFKDVNQTIDHGPWASFWGESGA 302
QY 348 YGGGAPLLSDTFPAAGFMWLDKGLSARMGIEVVMROVFFGAGNYHLVDE-NPDPLPDYWL 406
DB 303 YNSGGRHVSDTFIDSFWYLDQLGMSARHNTKYVCHQTLVG-GFYGLLEKGFVPPNDYIS 361
QY 407 SLLFKKLVGTRVMASVQSSKRRKRLVYLHCTNTDNPYKEGDLTYAINLHNVTYKL-- 464
DB 362 ALLMHRMLGKGVLAQVTDGP--QLRYAHCK-----GRAGVTLLILNLSNGDFTVS 413
QY 465 -----RLPYPS---NKQVDKYLLRP--LGRHG--LISKSVOL 495
DB 414 VSGNINVLNAESRKKSLLDTLKRPFWSIGSKASDGYLNREEYHLTPENGVLRSKTYVL 473
QY 496 NGULTKMWDDQTLPELMEKPLRPGSSSLGPAFSSYFFVIRAKYAAAC 542
DB 474 NGSKLEPTATGDI PSL-EPVLRSVNSPLNVLPLSMSTIVLPNFDASAC 520

```

Search completed: March 23, 2005, 11:29:35
 Job time : 111.5 secs

```

Db 78 FLSQLDPSIIHD-GWLDPLSKRLVTLARGSPAFIRGCKRTDPLQONLRNPAKSRG 136
Qy 112 FEERSYQVNOVODI-----CKYGSIPRVEEKLRLMEYQOQL-LREHYOK 158
Db 137 GGPDPYLYKNYDDIIRSDVALDKQCKCIAQ-HPVMLELQREKAQMLVLLKQF-- 193
Qy 159 KEKNSYSSRSVAVLYTFANCSCGLDLIFGLNALLRTADLQNMSSNAQLLLDYCSSKGVNI 218
Db 194 ---SNTYS-----NULL----- 202
Qy 219 SWELGNEPNSFLKKADIFINGSQLGEDFIQLHKLRLK-STFKNAKLYGPDVGQPRRKTA 277
Db 203 -----TEPNNYRTMHGRAVNGSQLGKDYIQLKSLQPIRISFASLYGNIGRPRKNVIA 257
Qy 278 MLKSLFKAGGEVYDSVTMHYYLNGRTATREDFLNDVLDIFISSVQKYFOVYESRPRK 337
Db 258 LIDGFKKVAGSTVDATVQHCYIDGKRVKMDFLKTRLLDTSDDIRKIQKVNTYTPGK 317
Qy 338 KYMLGETSSAYGGAAPLSDTFAGFPMWLDKGLSARMGIEVVMROVFGAGNYHLVDEN 397
Db 318 KIMLEGVVTTASAGTNNLSDSYAAGFLWINTLGMLANOGIDVYIRHSFDDHGNHLVDQN 377
Qy 398 FPPLEPDYMLSLFKKLVGTRKVLMAVQSKRR-----KLRVYLHCTNTDNPRYKEG 448
Db 378 FNPPLDYMLSLYKRLIGPKVLAVHVAQLQKRPGRVIRDKLRIRYAHCTNNHNNHYVRG 437
Qy 449 DLTVAIINAHNTKYRLPYPPSNKQVDKYLRLPGPHGLSKSVQNLGTLKMYDDOTL 508
Db 438 STLFPIINHRSRKIKLGLTRDKLVHGYLLQPYQOEGLSKSVQNLQGPLVWDDGTL 497
Qy 509 PRLMEKPLRPGSSLGPAFSYSPFVIRNAKVAAC 542
Db 498 PELKRPRLRAGRTLVIPVYTMGFVVKNNALAC 531

```

RESULT 13

```

O9HB39 PRELIMINARY; PRT; 480 AA.
ID O9HB39
AC O9HB39;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Heparanase-like protein HPAA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
OX NCBI_TaxID=9606;
RN RP SEQUENCE FROM N.A.
RX SEQUENCE-20483645; PubMed=11027606; DOI=10.1006/dbrc.2000.3586;
RA McKenzia E., Tyson K., Stamps A., Smith P., Turner P., Barry R.,
RA Hiccock M., Patel S., Barry B., Stuberfield C., Tetreu J., Page M.;
RT "Cloning and expression profiling of Hpa2, a novel mammalian
RL heparanase family member.";
RL Biochem. Biophys. Res. Commun. 276:1170-1177(2000).
RN RP SEQUENCE FROM N.A.
RA McKenzia E.A., Tyson K., Stamps A.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; A282885; AAG23421.1; -.
DR PIR; JC7506; JC7506.
DR InterPro; IPR005199; Glyco_hydro_79N.
DR Pfam; PR03662; Glyco_hydro_79n; I.
SQ SEQUENCE 480 AA; 5390 MW; F75F89F67AC1FF83 CRC64;

```

Query Match 31.5%; Score 893.5; DB 2; Length 480;
 Best Local Similarity 35.8%; Pred. No. 5,1e-59;
 Matches 201; Conservative 75; Mismatches 146; Indels 139; Gaps 9;

```

Qy 20 PLGPIUSPGL-----PRPA-----QAQVVDLDFIQEQLHVSRS 55
Db 18 PACIAPGALYALLLHLISLSSQAGDRRLPVDRAAGLKEKTLILLDVSTKKNRVTNEN 77

```

```

Qy 56 FLSVTIDANLATDPRFLILGSPKLTTLARGSPAVLRFEGTKTDFLIF-----DPKKEST 111
Db 78 FLSQLDPSIIHD-GWLDPLSKRLVTLARGSPAFIRGCKRTDPLQONLRNPAKSR- 135
Qy 112 FEERSYQVNOVODICKYGSIPRVEEKLRLMEYQOQLLREHYOKKNSYSSSDV 171
Db 136 -----GGGPPD-----YYLKNE----- 148
Qy 172 VLYTFANCSCGLDLIFGLNALLRTADLQNMSSNAQLLLDYCSSKGVNISWELGNEPNSFLK 231
Db 149 -----DEPNNYRT 156
Qy 232 KADIFINGSQLGEDFIQLHKLRLK-STFKNAKLYGPDVGQPRRKTA MLKSLFKAGGEVY 290
Db 157 MHGRAVNGSQLGKDYIQLKSLQPIRISFASLYGNIGRPRKNVIALIDGFKKVAGSTV 216
Qy 291 DSVTMHYYLNGRTATREDFLNDVLDIFISSVQKYFOVYESRPRK KYMLGETSSAYG 350
Db 217 DATVQHCYIDGRVVKVMDFLKTRLLDTSDDIRKIQKVNTYTPGK KIMLEGVVTTASG 276
Qy 351 GAPLSDTFAGFPMWLDKGLSARMGIEVVMROVFGAGNYHLVDENPPLDPYMLSLF 410
Db 277 GTNNLSDSYAAGFLWINTLGMLANOGIDVYIRHSFDDHGNHLVDQNPPLDPYMLSLY 336
Qy 411 KKLVGTRKVLMAVQSKRR-----KLRVYLHCTNTDNPRYKEGDLTLVAIINAHNT 461
Db 337 KKLIGPKVLAVHVAQLQKRPGRVIRDKLRIRYAHCTNNHNNHYVRG STLFIINHRSR 396
Qy 462 KYRLPYPPSNKQVDKYLRLPGPHGLSKSVQNLGTLKMYDDOTL PRLMEKPLRPGSS 521
Db 397 KKIKLAGTRDKLVHGYLLQPYQOEGLSKSVQNLQGPLVWDDGTL PELKRPRLRAGRT 456
Qy 522 LGLPAFSYSPFVIRNAKVAAC 542
Db 457 LVIPVYTMGFVVKNNALAC 477

```

RESULT 14

```

O8T108 PRELIMINARY; PRT; 515 AA.
ID O8T108
AC O8T108;
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE Heparanase-like protein.
OC Bombbyx mori (Silk moth).
OC Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Dictysta; Bombycoidea;
OC Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN RP SEQUENCE FROM N.A.
RC STRAIN=D50; TISSUE=Posterior silk gland;
RA Koike Y., Mita K., Suzuki M.G., Maeda S., Abe H., Osoegawa K.,
RA deJong P.J., Shimada T.;
RT "Genomic sequence of a 320-kb segment of the Z chromosome of Bombyx
RT mori containing a ketilin ortholog.";
RL Mol. Genet. Genomics 269:137-149(2003).
DR EMBL; AB079860; BAB85191.1; -.
DR EMBL; AB090307; BAC10612.1; -.
DR InterPro; IPR005199; Glyco_hydro_79N.
DR Pfam; PR03662; Glyco_hydro_79n; I.
SQ SEQUENCE 515 AA; 59769 MW; FB8100AB86EDDADB CRC64;

```

Query Match 24.6%; Score 699; DB 2; Length 515;
 Best Local Similarity 35.2%; Pred. No. 3,1e-44;
 Matches 184; Conservative 82; Mismatches 182; Indels 74; Gaps 18;

```

Qy 46 QEPHLVVSFLSVITDANLATDPRFLILGSPKLTTLARGSPAVLRFEGTKTDFLIFD 105
Db 42 QEDIKLISEDPLSGID-TIETENYRNINYSPTRLRELAALSPARLRGSGTWSERLIF- 99

```

```

QY 20 PLGLPSFAL-----PRPA-----QAQDVVDLDFTOEBPLHVS 55
| : : : : :
DB 18 PRACIAPGALYLLALLHLSSQAQDRRLPVBRAGLKEKTLILDVSTKNVPTVNNEN 77
QY 56 FLSTVTDANLATDPRFLIILGSPKLTARGLSPAYLRFGGKTDFLI-----DPKKEST 111
| : : : : :
DB 78 FLSTLODPSI IHD-GWLDFLSSKRLVTLARGLSPAYLRFGGKTDFLIQONTLNPAKSRG 136
QY 112 FEERSYQOVNODI-----CKYGSIPPEVEEKLRLWEPYQOL-LYREHYOK 158
| : : : : :
DB 137 GGGPDYLLKKNYEDIVRSVDALDKOKGCKIAQ-HPDWMELQREKAAQHLVLLKQFEN 195
QY 159 KERNSTYSSSVVDVLYTFANCGLDILFGNALLRTPADLQWNSNAQLLDYSSSKGYNI 218
| : : : : :
DB 196 TYSNLLTLFARSLDKLYNSADCSGLHLIFALNALRRPNNSWSSALSLTKYSASKYNI 255
QY 219 SWEIENEPNPFKADIFINGSQGEDFIQHLKLR-STFKNAKLYGPDVQPRKRTAK 277
| : : : : :
DB 256 SWEIENEPNNYRTMHGRAVNGSLQGRDYIQLKSLQPIRISYRASLYGPRIGRKNVIA 315
QY 278 MLKSPFKAGGEVIDSVTHHYLLNGRTATREDPLNDVLDIFISSYQVQVVESTRPGK 337
| : : : : :
DB 316 LLDGFKAAGSTYDVAITWQHCYIDGRVVKVMDFLKTRLDLSDQIRKLOKVNTYTPGK 375
QY 338 KVMIGETSSAYGGAPLSDTPFAAGFMWLDKGLSARMGIEVVRQVFGAGNYHLVDEN 397
| : : : : :
DB 376 KTWLBEVVTTSAGGTNNLSDSYAAGFLMNTLTMLANOGIDVYIRHSFHDGYNHLVDQN 435
QY 398 FDBLPYVMSILFKLVGTKVLMASVQSGKRR-----KRVYLHCNTNTPRYKEG 448
| : : : : :
DB 436 FDBLPYVMSILFKLVGTKVLMASVQSGKRR-----KRVYLHCNTNTPRYKEG 495
QY 449 DTLVYINLHNVYKYLRLPYFSGNKQVDKYLRLPGHGLSKSVOLNGLTKMVDQTL 508
| : : : : :
DB 496 SITLFIINLHRSKKIKLACTLRDKLVHQYLLQPYQEGKSKSVOLNGLTKMVDQTL 555
QY 509 PPLMEKPLRPGSSGLPAPSYSPFVIRNAKVAAC 542
| : : : : :
DB 556 PELKPRPLRAGRTLVIPVTMGFFVNVNALAC 589

RESULT 11
Q9HWQ1 PRELIMINARY; PRT; 548 AA.
ID 08WMQ1
AC 08WMQ1;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Hepatanase 3.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Legoux P., Legoux R., O'Brien D., Salome M.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Pessague Saïontas B.J.O.P.S.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: A0299720; CAC82492.1; -
DR Pfam: PF03662; Glyco_hydro_79n; 1.
SQ SEQUENCE 548 AA; 61771 MW; B8986303FC73A60A CRC64;

Query Match 35.6%; Score 1011.5; DB 2; Length 548;
Best Local Similarity 41.6%; Pred. No. 7.1e-68;
Matches 223; Conservative 80; Mismatches 180; Indels 53; Gaps 9;

QY 8 ALPPLMLLLGLGLPSFAL-----PRPA-----QAQDVVDLDF 43
| : : : : :

```

```

DB 6 APPPEAMLSNSRPPACIAPGALYLLALLHLSSQAQDRRLPVBRAGLKEKTLILDV 65
QY 44 FTOEPLHVSPEFLSTVTDANLATDPRFLIILGSPKLTARGLSPAYLRFGGKTDFLI 103
| : : : : :
DB 66 STKNPRTVNTNENLISQLODPSI IHD-GWLDFLSSKRLVTLARGLSPAYLRFGGKTDFLI 124
QY 104 F-----DPKKESTFEERSYQOVNODI-----CKYGSIPPEVEEKLRLWEPYQ 147
| : : : : :
DB 125 FQNLRRNPAKSRGGPGPDYLLKKNYEDIVRSVDALDKOKGCKIAQ-HPDWMELQREKAAQ 183
QY 148 EQL-LYREHYOKKERNSTYSSSVVDVLYTFANCGLDILFGNALLRTPADLQWNSNAQL 206
| : : : : :
DB 184 MHLVLEKQPSNTYSLIILFARSLDKLYNSADCSGLHLIFALNALRRPNNSWSSALS 243
QY 207 LLDYSSSKGYNISWEIENEPNPFKADIFINGSQGEDFIQHLKLR-STFKNAKLYG 265
| : : : : :
DB 244 LTKYSASKYKYNISWEIENEPNNYRTMHGRAVNGSLQGRDYIQLKSLQPIRISYRASLY 303
QY 266 PDVQPRKRTAKMLKSLFKAGGEVIDSVTHHYLLNGRTATREDPLNDVLDIFISSYQK 325
| : : : : :
DB 304 PNIGRPRKNVIALLDGFKAAGSTYDVAITWQHCYIDGRVVKVMDFLKTRLDLSDQIRK 363
QY 326 VPQVVESTRPGKVMIGETSSAYGGAPLSDTPFAAGFMWLDKGLSARMGIEVVRQV 385
| : : : : :
DB 364 IQKVNTYTPGKKTWLEGVVTSAGGTNNLSDSYAAGFLMNTLTMLANOGIDVYIRHSF 423
QY 386 FGAAGNHLVDENEDPLDYWLSILFKLVGTKVLMASVQSGKRR-----KRVYLH 436
| : : : : :
DB 424 FHDGYNHLVDQNPLPDYWLSILFKLVGTKVLMASVQSGKRR-----KRVYLH 483
QY 437 CTNTDNPYKEGDLTLYAINLHNVYKYLRLPYFSGNKQVDKYLRLPGHGLSKS 492
| : : : : :
DB 484 CTNNHNNHNVYKSGITLFIINLHRSKKIKLACTLRDKLVHQYLLQPYQEGKSKST 539

RESULT 12
Q9HB38 PRELIMINARY; PRT; 534 AA.
ID 09HB38
AC 09HB38;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Hepatanase-like protein Hpa2b.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20483645; PubMed=11027606; DOI=10.1006/bbrc.2000.3586;
RA McKenzie E., Tyson K., Stamps A., Smith P., Turner P., Barry R.,
RA Hircok M., Patel S., Barry E., Stuberfield C., Terrett J.,
RT "Cloning and expression profiling of Hpa2, a novel mammalian
hepatanase family member.";
RN [2]
RP Biochem. Biophys. Res. Commun. 276:1170-1177(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA McKenzie E.A., Tyson K., Stamps A.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF282886; AAC23422.1; -
DR InterPro: IPR005199; Glyco_hydro_79n.
DR Pfam: PF03662; Glyco_hydro_79n; 1.
SQ SEQUENCE 534 AA; 60063 MW; C3DE5E900CB338C4 CRC64;

Query Match 32.8%; Score 932.5; DB 2; Length 534;
Best Local Similarity 37.6%; Pred. No. 6.6e-63;
Matches 216; Conservative 79; Mismatches 168; Indels 111; Gaps 12;

QY 20 PLGLPSFAL-----PRPA-----QAQDVVDLDFTOEBPLHVS 55
| : : : : :
DB 18 PRACIAPGALYLLALLHLSSQAQDRRLPVBRAGLKEKTLILDVSTKNVPTVNNEN 77
QY 56 FLSTVTDANLATDPRFLIILGSPKLTARGLSPAYLRFGGKTDFLI-----DPKKEST 111
| : : : : :

```


	Matches	321;	Conservative	86;	Mismatches	114;	Indels	11;	Gaps	3
Qy	13	LMLLLGLPGLSGALPRPAOADVDDLFDTGEPRLHVBSPFSTVITDANLATDPFL	72							
Dd	2	LVLILLVLLLAVP-----RTMELQGLRPEIPAVSAPAFSLTLDASLADPPFV	52							
Qy	73	IILDSPKLRITARGLSPAYLARFGGKTDFLLFDPKKESTFEERSYWGOSQVNDICKYSI	132							
Dd	53	ALLHPKLHTLASLSGFSGLFFGGTIDPFLFNPKDSITWEKKULSEQA-KDVEAMPSS	111							
Qy	133	PPDVEEKLRLMPYOELLREHYOKKPKNSYSHSVDIVLYTPANCSDLIFGLNALL	192							
Dd	112	FAVVPKLLLTQWPIQEKKLLAEHSKKCKNTTITRSTLIDIHTFASSGFRLVFGLNALL	171							
Qy	193	RTADLOWNSSAOLLIDYCCKNYINSMELNEPMSFLKAKDIFINGSQLGEDFIQLHKL	252							
Dd	172	RRAQLQMDSSNAKULLGYCAQRSNVINSMELNEPMSFKKSGICIDGFOLGADFPHLNQL	231							
Qy	253	L-RKSTFPMNAKLYGPDVGPARRKTAAMKSFLEKAGGEVIDSWTHHHYLNGRATREDPL	311							
Dd	232	LSQHPIVRHALYLGIDVGQRPKRKHQHLNRSTMKSCKAIIDVTWHYYLVNGRSAIREDFL	291							
Qy	312	NPDVLDFISSVQKVFOVESTREPCKRWALGETSSAYCGGAPLLSDTFPAAGFMWLDKGL	371							
Dd	292	SPEVLDSFATRIHDVLGIIEATVPGKKWALGETSGAYGGAPQLSNTYVAGFMWLDKGL	351							
Qy	372	SABRGIEVVMQOVFEGAGNYHJVENPDLTDYMILLFKKLGVKXVLMASVQSKRRKL	431							
Dd	352	AARGGIDVMMQVSEGAOSYHVLVDGFFPLPDMYSLTLKYRLVGRVLOASVEQADARRP	411							
Qy	432	RVYLACTNTDNPRYKESGDTLTYAINLNVTXYRLPYEPSNKOVDKYLLRPYPGLGLSK	491							
Dd	412	RVYLACTNPBRPKRKREGDVTLPALNLNVTSQIQPKQLWSKSDVQYLLPHGXDSLIR	471							
Qy	492	SVQUNGTLTKAVDDQTLPPLMEKPLRPSSSLGPLPFSSFYIRAKYAACI	543							
Dd	472	EVOJNGRLLOWVDETTLPALHEMALPASTLGLPAFSYGFIYIRAKAIACI	523							
RESULT 9										
O9HB37										
ID	O9HB37	PRELIMINARY;	PRT:	592	AA.					
AC	O9HB37;									
DT	01-MAR-2001 (TREMBLrel. 16,	Created)								
DT	01-MAR-2001 (TREMBLrel. 16,	Last sequence update)								
DE	01-JUN-2003 (TREMBLrel. 24,	Last annotation update)								
OS	Heparanase-like protein HPA2C.									
OC	Homo sapiens (Human).									
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.									
OX	NCBI_TaxId=9606;	[1]								
RN										
RP	SEQUENCE FROM N.A.									
RA	McKenzie E.A., Tyson K., Stamps A.;									
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.									
DR	EMBL; AF282887; AAC23423.1; -									
DR	GO; GO:0005622; C:intracellular; TAS.									
DR	GO; GO:0030305; F:heparanase activity; TAS.									
DR	InterPro: IPR005199; Glyco_hydro_79N.									
PC	pfam; PF03662; Glyco_hydro_79n; I									
QC	SEQUENCE 592 AA; 66580 MW; 95C384AD9A74258B CRC64;									

```

0Y 20 PLG6LSRGL-----PRP-----QADVVLDLDFEOPHLVPS 55
Db 18 PPA1LPAAGALYALLLHLSLSQAGDRRLPVPDRAGLKEKLLLDVSTKOPVTYVEN 77
0Y 56 FLAVTIDANLAPREFLLILGSPKRLTARGLSPAAYLRFEGKTDFLF---DPKEST 111
Db 78 FLBLQDPSI:IH-GWLDPLSKRLVTLARGLSPAFLRFGGRTDFLOFQULRNPAPSRG 133
0Y 112 PEERSYQOVNODI-----CKYGIIPDVEBKRLMPYEOU-LREHYOK 155
Db 137 GPGPDYLLKNYEDDIVRSDVALDKQKCKIAQ-HPDWLELQREKAQMHVLJKEOPSN 199
0Y 159 KFNKSTYSRSYVULYTPANCGLDILFETMLLFTALDLQWNSNAOULLDYCSSKXNI 211
Db 196 TYNLLITARSIDLKLNFADECGLHLEFNLALRPNPNNSWSSALSLLKYSASKYNI 255
0Y 219 SWE1GNEPNSFLKKADIFINGSQLGDFYQLHKLRLK-STFNALYLPDVGOPRKTAK 277
Db 256 SWE1GNEPNNRITMHGRAVNSQLGXDYIQLKSLLOPIRITSRLBYEPNIGRPKNVIA 311
0Y 278 MLKSLKAGGEVIDSYTMHYYLNGRTAREDFLNPVDLDFISSQVQFOVESRTRPK 333
Db 316 LLDGFMKVAGSTVDAVTWQHICYIDGCVVMKMFLLKTRLLDPLSDQIRIKQKVNTYTPGK 377
0Y 338 KWLNGEISSAYGGGALLDPTFAAGFMULDGLSARNGIEVYMQVFFGAGNHLVDEN 399
Db 376 KIMLEGVVTSAGGTNNLSDSYAAGFLMNTGLMLANGOIDIVYIHSFDFGHYHNLVDON 433
0Y 398 FDP1PDVWLILPFKKLVGTRKVLMAASVOGSKR-----KLRYLHCTNTDNPYKEG 444
Db 436 FNP1PDVWLSLYKRLIGPKVLAVNAGIQRPGRGVIRYDGLRIYALCTHHNNHNVYRG 499
0Y 449 DUTLYAINHNVTXYLRLPYPPSNKQVDKYLLRLPGRHGLSKSYVOLNGLTLKNVDOTL 500
Db 496 SITFLIINIRSKRKIKLGLTRDKLVHGYLLQPIYGOEGLSKSYVOLNGOPLVWVDGTL 555
0Y 509 PPLMEKRLRGSSLGLPAPSYSPFYIRNAKVAAC 542
Db 556 PELKPRRLRAGRTLVPVYMGFVAVKNVNAALC 589

```

```

RESULT 10
Q8WMQ2
ID Q8WMQ2 PRELIMINARY; PRT; 592 AA.
AC Q8WMQ2;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)
DE Hepatranase 2.
OS Homo sapiens (Human) .
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Legoux P., Legoux R., O'Brien D., Salome M.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Pesseque Safontas B., J.O.P.S.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ299719; CAC82491.1; -
DR Genew; HGNC:18374; HPS82.
DR Pfam; PF00662; Glyco_hydro_79a; 1.
SQ SEQUENCE 592 AA; 66520 MW; 9478841FEACD58B CRC64;

Query Match: 40.2%; Score 1142.5; DB 2; Length 592;
Best Local Similarity 43.2%; Pred. No. 9.4e-78;
Matches 248; Conservative 83; Mismatches 190; Indels 53; Gaps 9

```



```

RESULT 3
Q9MYXO PRELIMINARY; PRT; 545 AA.
ID Q9MYXO
AC Q9MYXO;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hepatanase.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=21176669; PubMed=1127877;
RA Kizaki K., Nakano H., Nakano H., Takahashi T., Imai K., Hashizume K.;
RT "Expression of heparanase mRNA in bovine placenta during gestation.";
RL Reproduction 121:573-580(2001).
DR EMBL; AF281160; AAF87301.2; -
DR InterPro: IPR005199; Glyco_hydro_79N.
DR Pfam: PF03662; Glyco_hydro_79n; I.
SQ SEQUENCE 545 AA; 61076 MW; FAC4BDFPD855B933 CRC64;

Query Match 80.4%; Score 2285; DB 2; Length 545;
Best Local Similarity 80.0%; Pred. No. 2, 2e-164;
Matches 436; Conservative 34; Mismatches 73; Indels 2; Gaps 1;

QY 1 MLIRSRPALPPLMLL-LIGPLGIPSPGALPPPAQADVDLDFTQEPHLVSPSPIS 58
DB 1 MACCRKRGRLPPLLLPLGLPGPCSPGPAADADAELEFTEPRLHVSAPFIS 60
QY 59 VITDANLADPRLLILGSPKRTLRGSPALRGSTDTPLIDPKKESFEERSW 118
DB 61 FTIDANLADPRFTPLGSSKRTLRGALPAYLRGNGGDLPLDPKKEPFEERSW 120
QY 119 QSQVNODICKYGSIPPDVEKRLLEMPYQOLLRREHYOKKFNSTYSSSSVDLYTFAN 178
DB 121 LQSNODICKSGSIPSDVEKRLLEMPYQOLLRREHYOKKFNSTYSSSSVDLYTFAN 180
QY 179 CSGLDLIFGLNALLRTADLQWSSNAQLLDYCSSKGVNLSWELGNEPNSFLKADI 238
DB 181 CSGLNLIIFGVNALLRTDMDSSNAQLLDYCSSKGVNLSWELGNEPNSFLKADI 240
QY 239 GSGLDGDFLOHLLKSTFKRKAKLGVDPVGGPRRTAKMLKSLKAGGVLDISVTHNY 298
DB 241 GRLGDFLEFRKLLGSAFKAKMLYGPDIQGPRTVTKMLKSLKAGGVLDISVTHNY 300
QY 299 YLNGRATEDPLNPVDLIFISSVQKVPQVVESTRPGKKVMGGETSSAYAGGAPLSDT 358
DB 301 YVNGRATEDPLNPVDLIFISSVQKTLRIVEKIRPKKVMGGETSSAYAGGAPLSDT 360
QY 359 FAAGFMWLDKLGASNGIEVVMROVFFGAGNYHVDENFDPLDYWLSLFFKLVGTGY 418
DB 361 FAAGFMWLDKLGASNGIEVVMROVFFGAGNYHVDENFDPLDYWLSLFFKLVGTGY 420
QY 419 LMASVQSGSRKRLRYLHCTNTNPRYKSGDLTLVAINLHNVKYLRLPFPNKQVDKY 478
DB 421 LMASVQSGSRKRLRYLHCTNTNPRYKSGDLTLVAINLHNVKYLRLPFPNKQVDKY 480
QY 479 LRLPLPHGLKSKVOLNGLTLKAVDDQTLPLMEKPLRGSSGLGPAFYSFVIRNAK 538
DB 481 LRLPSGDLGSSVOLNGLTLKAVDDQTLPLMEKPLRGSSGLGPAFYSFVIRNAK 540
QY 539 VAACT 543
DB 541 VAACT 545

```

```

AC Q9K3K3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Heparanase (Mus musculus 0 day neonate thymus cDNA, RIKEN full-length
DE enriched library, clone:A43010LM04 product:heparanase, full insert
DE sequence).
GN Name=Hpa; Synonyms=Hpa, Hsps;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB;
RA Miao H.-O., Navarro E., Patel S., Sargent D., Koo H., Wan H.,
RA Plata A., Zhou Q., Ludwig D., Bohlen P., Kusie P.;
RT "Cloning, expression, and purification of mouse heparanase.";
RL Protein Expr. Purif. 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sami N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Oneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuda S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Horii F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,

```

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schermer A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences,"
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RA Strausberg R.;
 RL Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Vlodavsky I., Friedman Y., Elkin M., Aingorn H., Atzmon R.,
 RA Ishai-Michaeli R., Bitan M., Pappo O., Peretz T., Michael I.,
 RA Spector L., Pecker I.;
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF161514; AAD45379.1; -;
 DR EMBL; AF152376; AAD45669.1; -;
 DR EMBL; AF155510; AAD54941.1; -;
 DR EMBL; BC051321; AAH51321.1; -;
 DR EMBL; AF144325; AAD41342.1; -;
 DR GO; GO:0004566; F-beta-glucuronidase activity; TAS.
 DR GO; GO:0006029; P:proteoglycan metabolism; TAS.
 DR InterPro; IPR005199; Glyco_hydro_79N.
 DR Pfam; PF03662; Glyco_hydro_79n; 1.
 DR CHAIN 158 543 heparanase.
 SQ SEQUENCE 543 AA; 61176 MW; AD262EC267334AB2 CRC64;

Query Match 100.0%; Score 2841; DB 2; Length 543;
 Best Local Similarity 100.0%; Pred. No. 1.6e-206;
 Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLRRKPPALPPMLLLGLPLSPGALPPPAQADVDLDFPTQEPHLVSPFLST 60
 DB 1 MLRRKPPALPPMLLLGLPLSPGALPPPAQADVDLDFPTQEPHLVSPFLST 60
 QY 61 IDANLATDPRFILLGSPKRLTARGLSPAYLRFGGTKTDFLIPDKKESTFEERSYM 120
 DB 61 IDANLATDPRFILLGSPKRLTARGLSPAYLRFGGTKTDFLIPDKKESTFEERSYM 120
 QY 121 QVNVODICKYGSIPPDVEEKLRLMPYQEOQLLREHYQKKFKNSTYSSSDVLYTFPANC 180
 DB 121 QVNVODICKYGSIPPDVEEKLRLMPYQEOQLLREHYQKKFKNSTYSSSDVLYTFPANC 180
 QY 181 GIDLIFGLNALRLTADLQWSSNAQLLDYCSSKGYNISWELGNEPNSFLKKADIFIN 240
 DB 181 GIDLIFGLNALRLTADLQWSSNAQLLDYCSSKGYNISWELGNEPNSFLKKADIFIN 240
 QY 241 QUGEDFIQLHKLRLKSTFKNAKLYGPDVGQPRKRTAKMLKSLKAGGEVIDSVTMHHY 300
 DB 241 QUGEDFIQLHKLRLKSTFKNAKLYGPDVGQPRKRTAKMLKSLKAGGEVIDSVTMHHY 300
 QY 301 NGRTATREDFLNPDVDLDFISSVQKVFQVVESTRPCKKWLGETSSAYGGAPLLSDTP 360
 DB 301 NGRTATREDFLNPDVDLDFISSVQKVFQVVESTRPCKKWLGETSSAYGGAPLLSDTP 360
 QY 361 AGFMMLDKGLSARMGIEVVMROVFFGAGNYHLVDENPDLPIYMLSLFKLVGKTV 420
 DB 361 AGFMMLDKGLSARMGIEVVMROVFFGAGNYHLVDENPDLPIYMLSLFKLVGKTV 420
 QY 421 ASVQSGSKRRKRLRYVLIHCTNDNPRYKSGDLTYAIVLHNTYKRLPYPPSKNOVDKY 480
 DB 421 ASVQSGSKRRKRLRYVLIHCTNDNPRYKSGDLTYAIVLHNTYKRLPYPPSKNOVDKY 480
 QY 481 RPLGPHGLLSKSVQNLGLTLKAVDDQTLPLMEKPLRPSSSLGLPAFSYFFVIRNAKY 540
 DB 481 RPLGPHGLLSKSVQNLGLTLKAVDDQTLPLMEKPLRPSSSLGLPAFSYFFVIRNAKY 540
 QY 541 ACT 543
 DB 541 ACT 543

RESULT 2
 Q9ULJ39
 ID Q9ULJ39 PRELIMINARY; PRT; 545 AA.
 AC Q9ULJ39, 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Heparanase.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=20229546; PubMed=10764835; DOI=10.1093/glycob/10.5.467;
 RA Dempsey L.A., Plummer T.B., Coombes S.L., Platt J.L.;
 RT "Heparanase expression in invasive trophoblasts and acute vascular
 RT damage,"
 RL Glycobiology 10:467-475 (2000).
 DR EMBL; AF084467; AAD54516.1; -;
 DR Genew; HGNC:5164; HSE.
 DR InterPro; IPR005199; Glyco_hydro_79N.
 DR Pfam; PF03662; Glyco_hydro_79n; 1.
 SQ SEQUENCE 545 AA; 61417 MW; 67B80ACD73C5A9A1 CRC64;

Query Match 99.3%; Score 2820; DB 2; Length 545;
 Best Local Similarity 99.6%; Pred. No. 6.3e-205;
 Matches 543; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1 MLRRKPPALPPMLLLGLPLSPGALPPPAQADVDLDFPTQEPHLVSPFLST 58
 DB 1 MLRRKPPALPPMLLLGLPLSPGALPPPAQADVDLDFPTQEPHLVSPFLST 60
 QY 59 VTIDANLATDPRFILLGSPKRLTARGLSPAYLRFGGTKTDFLIPDKKESTFEERSYM 118
 DB 61 VTIDANLATDPRFILLGSPKRLTARGLSPAYLRFGGTKTDFLIPDKKESTFEERSYM 120
 QY 119 OSQVNVODICKYGSIPPDVEEKLRLMPYQEOQLLREHYQKKFKNSTYSSSDVLYTFPANC 178
 DB 121 OSQVNVODICKYGSIPPDVEEKLRLMPYQEOQLLREHYQKKFKNSTYSSSDVLYTFPANC 180
 QY 179 CSGDLIFGLNALRLTADLQWSSNAQLLDYCSSKGYNISWELGNEPNSFLKKADIFIN 238
 DB 181 CSGDLIFGLNALRLTADLQWSSNAQLLDYCSSKGYNISWELGNEPNSFLKKADIFIN 240
 QY 239 GSQUGEDFIQLHKLRLKSTFKNAKLYGPDVGQPRKRTAKMLKSLKAGGEVIDSVTMHHY 298
 DB 241 GSQUGEDFIQLHKLRLKSTFKNAKLYGPDVGQPRKRTAKMLKSLKAGGEVIDSVTMHHY 300
 QY 299 YINGRTATREDFLNPDVDLDFISSVQKVFQVVESTRPCKKWLGETSSAYGGAPLLSDTP 358
 DB 301 YINGRTATREDFLNPDVDLDFISSVQKVFQVVESTRPCKKWLGETSSAYGGAPLLSDTP 360
 QY 359 FAAGFMMLDKGLSARMGIEVVMROVFFGAGNYHLVDENPDLPIYMLSLFKLVGKTV 418
 DB 361 FAAGFMMLDKGLSARMGIEVVMROVFFGAGNYHLVDENPDLPIYMLSLFKLVGKTV 420
 QY 419 LMASVQSGSKRRKRLRYVLIHCTNDNPRYKSGDLTYAIVLHNTYKRLPYPPSKNOVDKY 478
 DB 421 LMASVQSGSKRRKRLRYVLIHCTNDNPRYKSGDLTYAIVLHNTYKRLPYPPSKNOVDKY 480
 QY 479 LIRPLGPHGLLSKSVQNLGLTLKAVDDQTLPLMEKPLRPSSSLGLPAFSYFFVIRNAKY 538
 DB 481 LIRPLGPHGLLSKSVQNLGLTLKAVDDQTLPLMEKPLRPSSSLGLPAFSYFFVIRNAKY 540
 QY 539 VAACT 543
 DB 541 VAACT 545

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2005, 11:18:07 ; Search time 109.5 Seconds
(without alignments)
2539.352 Million cell updates/sec

Title: SEQ2B
Perfect score: 2841
Sequence: 1 MLRSKPALPPLMLLGP.....LPFSYSPFVIRNAKVAACI 543

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2841	100.0	543	2	Q9Y251 homo sapien
2	2820	99.3	545	2	Q9UL39
3	2285	80.4	545	2	Q9MYX0
4	2153	75.8	535	2	Q8K3K3
5	2143	75.4	535	2	Q6Y2Z1
6	2138	75.3	536	2	Q71RP1
7	2126	74.8	536	2	Q9QZK8
8	1648.5	58.0	523	2	Q9QYK5
9	1150.5	40.5	592	2	Q9HB37
10	1142.5	40.2	592	2	Q8WMQ2
11	1011.5	35.6	548	2	Q8WMQ1
12	932.5	32.8	534	2	Q9HB38
13	893.5	31.5	480	2	Q9HB39
14	699	24.6	515	2	Q8T108
15	417	14.7	521	2	Q9SDA1
16	417	14.7	543	2	Q9FRI0
17	400	14.1	559	2	Q89F99
18	392.5	13.8	544	2	Q8H815
19	379	13.3	541	2	Q69116
20	377	13.3	527	2	Q9LRC8
21	365	12.8	537	2	Q70VJ3
22	363	12.8	536	2	Q9FPIK
23	353.5	12.4	516	2	Q9FPIK
24	353.5	12.4	539	2	Q8L608
25	350.5	12.3	529	2	Q6Z0E2
26	169.5	6.0	190	2	Q82604
27	156	5.5	935	2	Q9VFE7
28	141	5.0	559	2	Q7SFB0
29	138.5	4.9	463	2	Q63T97
30	126.5	4.5	493	2	Q9HK01
31	118.5	4.2	408	2	Q9HEZ1

32	118.5	4.2	408	2	Q9HEZ2	Q9he22 phanerochae
33	116.5	4.1	617	2	Q40996	Q40996 measles vir
34	116	4.1	390	2	Q8TPH7	Q8TPH7 methanosarc
35	114.5	4.0	1169	2	Q869K5	Q869K5 dictyostell
36	114	4.0	398	2	Q72RP7	Q72RP7 leptospira
37	114	4.0	398	2	Q8F410	Q8F410 leptospira
38	113.5	4.0	617	2	Q83295	Q83295 measles vir
39	113.5	4.0	1829	2	Q9KH44	Q9KH44 pantoea agg
40	112.5	4.0	617	2	Q83647	Q83647 measles vir
41	112	3.9	732	2	Q7MSD4	Q7MSD4 wolinnella s
42	112	3.9	2319	2	Q96U00	Q96U00 neurospora
43	111.5	3.9	617	2	Q40991	Q40991 measles vir
44	111.5	3.9	617	2	Q98VH5	Q98VH5 measles vir
45	111	3.9	670	2	Q9M090	Q9M090 arabidopsis

ALIGNMENTS

RESULT 1	ID	Q9Y251	PRELIMINARY;	PRT;	543 AA.
AC	Q9Y251	01-NOV-1999 (TrEMBLrel. 12, Created)			
DT	01-NOV-1999 (TrEMBLrel. 12, Last sequence update)				
DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)				
DE	Heparanase.				
GN	Name=HPA; Synonyms=HPSE;				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
NCBI	NCBI_TaxID=9606;				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Placenta;				
RX	MEDLINE=99321249; PubMed=10395326;				
RA	Medelt M.D., Freeman C., Handorf B.J., Baker R.T., Harris M.J., Parish C.R.;				
RT	"Cloning of mammalian heparanase, an important enzyme in tumor invasion and metastasis."				
RL	Nat. Med. 5:803-809(1999).				
RN	(2)				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Placenta;				
RX	MEDLINE=99335379; PubMed=10405343; DOI=10.1006/dbrc.1999.0962;				
RA	Kusile P.H., Holmes J.D., Ludwig D.L., Patel S., Navarro E.C., Sedden A.P., Giorgio N.A., Bohlen P.;				
RT	"Cloning and functional expression of a human heparanase gene."				
RL	Biochem. Biophys. Res. Commun. 261:183-187(1999).				
RN	(3)				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99377052; PubMed=10446189; DOI=10.1074/jbc.274.34.24153;				
RA	Toyoshima M., Nakajima M.;				
RT	"Human heparanase. Purification, characterization, cloning, and expression."				
RL	J. Biol. Chem. 274:24153-24160(1999).				
RN	(4)				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Pancreeas;				
RX	MEDLINE=2288257; PubMed=12477932; DOI=10.1073/pnas.242603899;				
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schaller G.D., Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diachenko L., Martinsina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo W.F., Caesavart I.L., Schetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Pangue C., Raha S.S., Loquellano N.A., Peters G.J., Abrazon R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale K.C., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.;				

This Page Blank (uspto)

A:Gene: ADE1; SPDB:SPBC405.01

A:Map position: 2

C:Superfamily: Saccharomyces cerevisiae ADE5 multifunctional protein; phosphoribosylamin

C:Keywords: cyclo-ligase; putine nucleotide biosynthesis

F/5-425/Domain: phosphoribosylamine-glycine homology <PGL>

F/439-767/Domain: phosphoribosylformylglycinamide cyclo-ligase homology <PFL>

Query Match 3.7%; Score 104.5; DB 1; Length 788;

Best Local Similarity 27.7%; Pred. No. 15;

Matches 70; Conservative 36; Mismatches 114; Indels 33; Gaps 11;

QY 297 HYLNGRTATRE--DFLNDPV--LDIFISSVQKVQVVEST--RPGKKWLGETSSAY--- 348

DB 424 HHALNPKRKTRREILTYENGSGVDNNGNEVQRKIDLVKSTRRPGADADIGCGJFDLQ 483

QY 349 -GGAPLL--SDTPAAGFMMLDKGLSAR--MGEEVVMRQVFGAGGYHLDENPDPL--P 402

DB 484 AGMNDPLVASATDGVGSKLLIALSLNKHDTVTGIDLVAMNV-----NDLVQGAEPFL 537

QY 403 DYMLSLFFKKLVGTKYLMAVSGSKRRKRLRVYLHCTNTDNPRYKEGDLTLVAINLHNVTX 462

DB 538 DYFAISGLDKVSTSVGEVVKCKQAGCALVGETSEMPGLYHGHYDANGTSVAVSR 597

QY 463 YRLPYFESNKQVQKYLRLPLGPHGLSKSVQNLGLTL--KNVD---DQTLPLMEKPL 516

DB 598 DDLPRKESFSGKDILL-----GLASGVHNSNGYSLVRKIYEYSDLEVTSCPMKDV 650

QY 517 RRGSSGLPAFST 529

DB 651 RLGDSTLILPTRIY 663

RESULT 15

F70411

adenylosuccinate synthetase - Aquifex aeolicus

C:Species: Aquifex aeolicus

C:Date: 08-May-1998 #revision 08-May-1998 #text_change 09-Jul-2004

C:Accession: F70411

R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Oy

V. Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70300; WUID:98196666; PMID:9537320

A:Accession: F70411

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Residues: 1-432 <AOF>

A:Cross-references: UNIPROT:O67321; GB:AB000733; NID:92983720; PIDN:AA07286.1; PID:9298

A:Experimental source: strain VFS

C:Genetics:

A:Gene: purA

C:Superfamily: adenylosuccinate synthase

Query Match 3.7%; Score 104; DB 2; Length 432;

Best Local Similarity 23.9%; Pred. No. 6.4;

Matches 96; Conservative 39; Mismatches 128; Indels 138; Gaps 22;

QY 15 LLLGLPLGSLPSPALRRPAQADVDL-----FTQEPHLVSPS 55

DB 51 IHLLEPTGLIHEHVKGVIAGM--VVDLEVLHKEVKNLEEKIYKERLFTSDRAHLVMPY 109

QY 56 FLSTVIDANLATDPRFLILGSPK--LRTIARGSPAYL--RFGGKTDFLIPPKKESTP 112

DB 110 H-----KLLDSLFKKKGIGTLRGIGPAYMFKY--RKGIISDLKDEKRF 154

QY 113 EERSYWGQVQVODICKYGIIPDVEEK-----LRLWPYQQLIREHYQKKFKNSTY 165

DB 155 ----YTLLEBNLDVFK-----NICEKVCEKFDLDINQIYEQL----RYFEFEKENV- 199

QY 166 SRSSVDVLYTFANCSGLDLFLGLNALRLTADL-----QMSNSNAQLLLDYCSSKGVNISWE 221

DB 200 ----VDLRFNTQKSGVLFEGAGTLLDVDMGTYPYVTSSNASAL-----GLSNG 246

QY 222 LGNEBNSFLKADIFING-----SGL-GEVYIQHLKLRKSTFKQAKLYG 265

DB 247 TGMPPKVF---SDAPFLGVAKAYTTRVGEGFPFTELKGEKEKREL-----GGEYG 295

QY 266 PVVGQPRR---KTAKMLKSLKAGGEVIDSVTWHHYLNGRTATREDPLN----- 313

DB 296 STTGRRPRRCGWLIDVALKRYAVQVNG-----LDGFVITKLDVLDTPDEVKVCVA 343

QY 314 -----DVLDIFISSVQKVQV--VESTPRGKKWLGETSSA 347

DB 344 YELDGEVIDYFPASYSBELRVKVPYKTLKG---WKKSTKGA 381

Search completed: March 23, 2005, 11:31:39

Job time : 30 secs


```

QY 164 TVSRSSVDVLYTFANCSGDLIFGL-NALLRTADLQMNSSN-----AQLLDLYC----- 211
DB 121 -----GIDIK---ADLSGALVLSLFOALEYTDIMWDPSPRWDGISGLDIADISITAO 182
QY 212 -----SSKGVNI-SMEL-GNEPNSFL--KKADIFINSQ----- 241
DB 183 RHEENGDDSSNEISGNGTVGNLGAARLRADMOCTDYLHSSKNDDBDVINGDPTQKWEWSR 242
QY 242 -----LGEDYIOLHKLRLKSTF-----KNAKLYGPDV 268
DB 243 YYAMRALPRLKAKGLIGEDY-----LNSDIFDGNFYVGSISTDQMLPRLRGAPDI 296
QY 269 GQPRKRTAMLSFLKAGEVI-----DSVTWHYYLNGRTATREDFLN 312
DB 297 SGVAHTTAKVTVSOL---GRVYETQVPAGPPIODLGDV-----SGTLHIREON 346
QY 313 PVLVDIFISSVQKVPQVSTRPGK---KVWL-----GENTS----- 345
DB 347 GGVQVEDINTASMPF---LTRPGQVRKLMGGRPOEWHVHGEGFSGSEASWGIANCW 402
QY 346 SAYGGAPPLSD-----TPAA-----GFWMLDKL-----GLSAR 374
DB 403 SLYCGA---LADENHQSALAGVGRDLSTVGAVAFDTHSHRTLDKETAVGKSLDGNFR 459
QY 375 MGI-----EVMNRQVFEAGNTHLVDENFDPDPLDYWLSLFFKLYGT---KVLMAVQGS 426
DB 460 LSYSKDFDELNRSRVFAG---YRFSEENMTMSEY-LDASDSBEMVATGNDKEMVATVNO 515
QY 427 KRRKLRVYLHCTNTNPRKEDDLTYAL-----NLHNVTK----- 462
DB 516 NFRDAGVSYYLVYTRTYVDRDEQTNVWLSHYFLGSIKRNMSISMTGYRYEYDNOADK 575
QY 463 --YLRLPYPSNKKQVDKYLRLPLGPHGLSKSVOLNGLTKLWDDPT 507
DB 576 GYIISLMSWGDSSSTISY-----NGNNGSGSDSSQVQ--YFSKVDAT 616

```

RESULT 10

hypothetical protein sir1617 - *Synechocystis* sp. (strain PCC 6803)

C.Species: *Synechocystis* sp.

A.Variety: PCC 6803

C.Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

C.Accession: S74760

R.Kaneko, T., Sato, S.; Kocani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O. K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A>Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

8.

A.Reference number: S74322; MUID:97061201; PMID:8905231

A.Accession: S74760

A>Status: preliminary

A.Molecule type: DNA

A.Residues: 1-411 <KAN>

A.Cross-references: UNIPROT:P72895; EMBL:D90901; GB:AB001339; NID:g1651897; PIDN:BA1691

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 3.8%; Score 108.5; DB 2; Length 411;

Best Local Similarity 21.4%; Pred. No. 2.7; Mismatches 94; Indels 93; Gaps 17;

Matches 66; Conservative 55; Mismatches 94; Indels 93; Gaps 17;

159 KFNSTYSSRSVDVLYTFANCSGDLIF--GMLALRTADLQMNSSNAQL----- 206

DB 151 EFRLLSPTRQIDID---FAGSTKLDLLASBENIDCVHLANPRVYTSNVAMGQTLMLAN 207

QY 207 LLDYSSSKG---YNI SMEL-----GNEPNSFLK-----ADIFINSQ LCE 244

DB 208 VIDVCIACKIDIPILYPSSEIYSGVAGTIAHDESTPALPRGYETKYLAELIL----- 260

QY 245 DVIQGLHLKRLKSFNKAUKYGPVQGPRRKRTAMLSFLKAGEVLDSTWHYYLNGRT 304

DB 261 DHCRRTRGLRICALRSSPVYSGWSMDP-----KFINPFRKASQGGKIYV--HHYING-- 311

```

QY 305 ATREDFLNPDV---LDIFISSVQKVPQVSTRPGKXVWLGETSAYGGAPPLSDPFA 360
DB 312 -----NPKLDLHLHIDLISSIVATL-----KSRFINLNI-----GTGQLSSTLK 351
QY 361 AGFWMLDKLGLSA-----RMGIEVVMRQVFPAGNTHLVDENFDDLPLDYWLSLFFKLYG 415
DB 352 IAEMLRDELSSSMIQIETVNTFVASIANVYGRAN-HVLD-----MEVYIFFE-QG 400
QY 416 TKVLMASV 423
DB 401 LKSLHLQI 408

```

RESULT 11

beta-xylosidase [imported] - *Caulobacter crescentus*

C.Species: *Caulobacter crescentus*

C.Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C.Accession: D87541

R.Nierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.I.

B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolton

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U S A. 98, 4136-4141, 2001

A>Title: Complete Genome Sequence of *Caulobacter crescentus*.

A.Reference number: A87249; MUID:21173698; PMID:11258647

A.Accession: D87541

A>Status: preliminary

A.Molecule type: DNA

A.Residues: 1-500 <STO>

A.Cross-references: UNIPROT:Q9A5U0; GB:AB005673; NID:g13423886; PIDN:AAK24328.1; GSPDB:IG

A:gene: CC2357

Query Match 3.8%; Score 107.5; DB 2; Length 500;

Best Local Similarity 25.7%; Pred. No. 4.4; Mismatches 80; Indels 47; Gaps 15;

Matches 56; Conservative 35; Mismatches 80; Indels 47; Gaps 15;

165 YSRSSVDVLYTFANCSGDLIFGLN---ALLRTAD---LQW--NSSNAQL-----LLD-- 209

DB 81 YDMTKIDQLYDALMLAKGIFELGFTPRAMKTSQDTIFYWKGNTHSHPLGFWRLIDAF 140

QY 210 -YCSSKGVNI---SW--ELGNEPN--SFLKADIFINSQGEYIOLHKLRLKSTFKN 260

DB 141 VHLRLARVGEVERRTWFFVWNEPNLDGWEKAD-----QAAVFEIYDV---TARA 188

QY 261 AKLYGPD--VQGPRTKRTAKMLKSLF---KAGEVIDSVTWHYYLNG---RTATREDPL 311

DB 189 IAIIDPSLRVGGPARAGAAWVEFLAHYKSSAANDFYTHHYGVGDGFLDEKGVDTKL 248

QY 312 NPDVLDIFISSVQKVPQVFE--STRPGKXVWLGETSAY 348

DB 249 SPSD-DAVVGDVVRVREQIEAGAPGLPLFYFTEWSTSY 285

RESULT 12

F85875

probable fibribral usher 23600 [imported] - *Escherichia coli* (strain O157:H7, substrain EI

C.Species: *Escherichia coli*

C.Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C.Accession: F85875

R.Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew,

Iller, L.; Grobleck, E.J.; Davis, N.W.; Lam, A.; Dimlatat, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A>Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.

A.Reference number: A85480; MUID:21074935; PMID:11206551

A.Accession: F85875

A>Status: preliminary

A.Molecule type: DNA

A.Residues: 1-879 <STO>

A.Cross-references: UNIPROT:O8XCP4; GB:AB005174; NID:g12516702; PIDN:AA657466.1; GSPDB:GT

A:Experimental source: strain O157:H7, substrain EDJ933

C:gene: 23600

C:Species: Saccharomyces cerevisiae
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
 C:Accession: S33961; S46140
 R:Doignon, F.; Bileau, N.; Crouzet, M.; Aigle, M.
 Yeast 9, 189-199, 1993
 A:Title: The complete sequence of a 19,482 bp segment located on the right arm of chromo-
 A:Reference number: S29348; MUID:93220397; PMID:846566
 A:Accession: S33961
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-688 <DOI>
 A:Cross-References: UNIPROT:P38338; EMBL:X70529; NID:G1907246; PID:CAA49923.1; PID:G296
 R:Aigle, M.; Bacle, M.C.; Barthe, C.; Bileau, N.; Crouzet, M.; Doignon, F.
 submitted to the Protein Sequence Database, August 1994
 A:Reference number: S45940
 A:Accession: S46140
 A:Molecule type: DNA
 A:Residues: 1-688 <AIG>
 A:Cross-References: EMBL:Z36128; NID:G536684; PID:CAA85222.1; PID:G536685; MIPS:YBR259W
 C:Genetics:
 A:Cross-References: SGD:S0000463
 A:Map position: 2R
 C:Superfamily: Saccharomyces cerevisiae hypothetical protein YBR259W

Query Match 3.9%; Score 111; DB 2; Length 688;
 Best Local Similarity 22.5%; Pred. No. 3.9;
 Matches 67; Conservative 45; Mismatches 94; Indels 92; Gaps 16;

```

QY 126 ICKYSGIPDVEKRLLEMPYQQLREHYQKKFKNSTYRSR-----VDLYT 175
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 164 MAEYSWKWDSDKQLOPFMEYEFKMLKECLVFYENFDLQKSDPLKELIIPWEKIVYV 223
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 176 PANGGLDIFGLNMLRTADLQWNSN-----AOILLD-----YCSKGY----- 216
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 224 -ANC--IDAFTEQVRIDAEILMTSKLVFSSISAVLRLLDQWMSAFRPGEBALV 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 217 -----NISWELGNEPNSFLKKA--DIF--INGSQLG--EDYIQLHLKLR----- 255
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 281 QDFAHRSIKWDSNKVESLIRALIFNDWFFYFNKEQVDTKADGIFPLRLKRFKEHIN 340
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 226 -----STFKK--AKLYGPDVGQPRRTAKMLKSFLLKAGEV-----IDSV 293
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 341 DVKDFHIQVIXKYLNSQFKNNYSTLMTSSKTQDRKSHNMPSSILLDGNKIGMHSPIDE- 399
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 224 TWHHYLLNG-----RTATREDPLNPDVLDFISVQVQVVESTR---PGKK 338
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 400 -YSHFIDNDEPLMRQKVYFKITNEQTPTPDASALFDS--HKIYAIISLRYLPEKR 454

```

RESULT 8
 T49648
 hypothetical protein B8B20.20 [imported] - Neurospora crassa
 C:Species: Neurospora crassa
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
 C:Accession: T49648
 R:Schulte, U.; Aigle, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
 submitted to the Protein Sequence Database, May 2000
 A:Reference number: Z25022
 A:Accession: T49648
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-2298 <SCH>
 A:Cross-References: UNIPROT:Q96U00; EMBL:AL355933; GSPDB:GN00116; NCSP:B8B20.20
 A:Experimental source: BAC clone B8B20; strain OR74A
 C:Genetics:
 A:Gene: NCSP:B8B20.20
 A:Map position: 6
 A:Introns: 426/3

Query Match 3.9%; Score 111; DB 2; Length 2298;
 Best Local Similarity 19.3%; Pred. No. 24;
 Matches 114; Conservative 79; Mismatches 190; Indels 208; Gaps 28;

```

QY 78 PGLRTANGLSPA-----YLFQGTGKTDFLIDPKKSTFEREYSQV-NODIC 127
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1447 PRVDIEIRLITPSNHNKACLINIRANQGLARLVVSNEGSAPFLITRRNVNQI LD 1506
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 128 KYGSIIPDVEEKL-----LEMPYQQLREHYQKKFKNSTYRSRSDVLYTPAN 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1507 QYMSAESDIEQOFALSAENNRSIDAWEELITRN-----KATLIDILHTSAR 1555
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 179 CSGDLIFGLNML-----LRTADLQ-----NNSNAQILLDCCSKGYNI 218
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1556 AS-LDVLQAKTLEAIIYTLNVTQLOKNCCTLHFQSPGFDGILVAL-----DTHANFL 1609
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 219 SW-ELGNEPNSFLKADIFINGSOQGEDYIQLHKLKSTFKNATLYGPDVGQPRRTAK 277
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1610 GMIFTSSEQYSSNSSADIDPROLEDAILLQELTTEPFMA-----RELLAL 1659
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 278 MLKSFLLKAGEVID-SVTWHHYLLNGRATREDPLNPDVLDFISVQVQVVESTRPG 336
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1660 PLKAITTEGKQTEQVACTEKTVTTLAKLAAR-----FLQ-ERVTVLPYFQPG 1706
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 337 K-----KWLGETSSAYGGAPLSDTFAAG-----PMWIDK 368
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1707 KYGLFPDMPKNGSGPERRWL-----PLFIATLVKNKVPDFKQIETNISTLWQS 1755
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 369 IGLSAR-MGIEVVMQVFFGAGNYHL--VDENFDPLPDYMLSL-LFKLLVG--TKVL--- 419
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1756 IIKPMPFLGYETLYLAELVQQRGLPFLAEDVSAGMTPPYNIHLDFSALHMRGLG 1815
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 420 -----MASVQSK-----RRKLRYVLIHCTNTDNPYKEDGLTYALNHNVTYL 464
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1816 ATPPAGVTSSASTGSSAQSIRQREFSH----- 1847
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 465 RLPEFNSKQDKYLLRLPG-----PHGLSSVOLNGITLGMVD-- 504
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1848 TLQALMTYIKDQLFLRLSLADLPASTTEHRDYMAFTGLIS-LISHGVGIIVVDSF 1906
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 505 -----DQTLPLMEKPLRPG-----SSIGLPAPFSYF-FVIRNAKVA 540
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1907 FLTPSDSYSPLODPLQHTAGIMAVGRLSEKDVPASQLFWYLENNKVA 1957

```

RESULT 9
 E91031
 probable outer membrane protein Ecs3221 [imported] - Escherichia coli (strain O157:H7, at-
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
 C:Accession: E91031
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
 Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shibata, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
 A:Reference number: A59629; MUID:21156231; PMID:11258796
 A:Accession: E91031
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-879 <HAY>
 A:Cross-References: UNIPROT:Q8XCP4; GB:BA000007; PID:BA816644.1; PID:G13362691; GSPDB:GR
 A:Experimental source: strain O157:H7, substrain RMD 0509952
 C:Genetics:
 A:Gene: Ecs3221

Query Match 3.9%; Score 109.5; DB 2; Length 879;
 Best Local Similarity 20.1%; Pred. No. 7.3;
 Matches 130; Conservative 68; Mismatches 208; Indels 241; Gaps 33;

```

QY 52 VSPSFLSVTIDANLATDPRFLILGSPRLRTLARGSLPAYRFGGTXDPLIDPKKST 111
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 20 MSGSYVNAWAEKEIQFDSRFLKGDTKI-DLKRRSSQGYVEPG--KYNLQVQLNKQPLT 76
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 112 PEERYSQVNOYDICKGSIIPDVEEKL-----RLLEMPYQQLREHYQKKFKNS 163
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 77 EBYDIWYASBNDASKTYACLTPELVAQGLKEVDANKMLQWIHQCKLPQGLE----- 130

```

Db 14 IRKMKIYNGKNEKDIKERLIKE-----LKEHVLVETEDGTYYTLKADEEEMNSKV 66
 QY 172 -----VLTFPANCGLDILFGNALLRTPADLQWSSNQMLLDYCSSKGYNISWEIGNEP 226
 Db 67 GALKKAIYKFAKPS-----KITDL-----SNPR-VLDLCSGKMGKNAIIMALHTNK 109
 QY 227 NS-----FL-----KKADIFINGSQLGEDIYQLHKLIRKSTF 258
 Db 110 NAEIMWEICEERVLFLTLFDLPYKEHEIITKDKREYFLN-KIGLEY-----KSDY 159
 QY 259 KNAKIYGPVQGPQRKRTAMKLSFLKAGGEVDSVTMHYIANGRTAT--REDFLNPDVL 316
 Db 160 DNINLY--VGDARKFIISKDKY-----NVFHDAPSPKDPITYTYDFL----- 202
 QY 317 DIFISSVOKFQVNVSTRPGKVMYLGETSAYGGAPLSDTFAGFMWLDKLGISARWG 376
 Db 203 -----KEIYKRMEDN--GVLI-----SYSSAIPFSSALVDGCFVISEKESVGRKKG 246
 QY 377 IEVNRQVFFGAGNYHLDENFD-----PLPDYMLSLFFKKLVGTVMASVOGSKRR 429
 Db 247 ILLAVKNPFRKPNRINEVDERVIATSVIALPYRDETLSTKDKITIEDREERREKLEKLI 306
 QY 430 KIRVYLHCNTNPNRYKEEDLTLTA--IYLAHVTKYLRPLPY 468
 Db 307 KIGKYLSTQIKKNIPEELIKIOKEDLNSSEIIRKMLKF 347

RESULT 5

T12094
 beta-fructofuranosidase (EC 3.2.1.26) - fava bean
 C:Species: Vicia faba (fava bean)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
 C:Accession: T12094
 R:Weber, H.; Borstjans, L.; Heim, U.; Buchner, P.; Wobus, U.
 Plant Cell 7, 1835-1846, 1995
 A:Title: Seed coat-associated invertases of Fava bean control both unloading and storage
 A:Reference number: Z17416; MUID:96093423; PMID:8535137
 A:Accession: T12094
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-575 <WEB>
 A:Cross-References: UNIPROT:Q43855; EMBL:Z25162; NID:g861154; PIDN:CAA84526.1; PID:g8611
 A:Experimental source: cv. Frito, seed coat
 C:Genetics:
 A:Gene: CM1N1
 C:Superfamily: beta-fructofuranosidase
 C:Keywords: cell wall; glycoprotein; glycosidase; hydrolase

Query Match 3.9%; Score 111.5; DB 2; Length 575;
 Best Local Similarity 21.4%; Pred. No. 2.7;
 Matches 72; Conservative 48; Mismatches 107; Indels 109; Gaps 19;

QY 46 QEPPLHVS-----PSFLSVITIDANLADPRFLILGSPKRTLARGLS-----P 89
 Db 228 KPHHSARKTGMECEDFYPVSLGKNGLD--LSMMGNVXHVKNLSLDIRREYTYTG 285
 QY 90 AYLR-----FGTKTDF-----LIPPKKESTFEEBSRW-----OSOVNQ 124
 Db 286 TYLQNDQDKTIPDKTSMDGKGLRYDGNFYASKSPDPTR-----NRRIIMGANEDDTGE 341
 QY 125 DICKYG-----SIPDV-----EEKLLEWYOEOLLR-----EHYOKKFNSTYSRSSV 170
 Db 342 DQVKGMAGIQAIPTVWLDSSRRQLR--QMPVEELRLRGQVEMNRRLKKGKY-----L 396
 QY 171 DVLVTFPANCGLDILFGNALLRTPADLQWSSNQMLLDYCSSKGYNISWEIGNEPFL 230
 Db 397 EVKGITASQADVVFSTSLDKAEAFDPNWEVAE--DLCAQSGSKVAGVG--PFGIL 451
 QY 231 KKADIFINGSQLGEDIYQL-----HKL-----RKSTFNKATLYGP-----DV 268
 Db 452 TLA-----SKLEBTTSTFFRVFRKANHGKILMCDASSSLNRELRYPSFAGFVNDL 505
 QY 269 GQPRKRTAMKLSFLKAGGEVIDSVTMHHYIANGRT 304

Db 506 GNNKKLSIRSL-----IDHSVESFVGSGKT 531

RESULT 6

T10666
 hypothetical protein F6E21.40 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
 C:Accession: T10666
 R:Bevan, M.; Lemnard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancroft
 submitted to the Protein Sequence Database, June 1999
 A:Reference number: Z16533
 A:Accession: T10666
 A:Molecule type: DNA
 A:Residues: 1-670 <BEV>
 A:Cross-References: UNIPROT:Q9M090; EMBL:AL049914; GSPDB:GN000662; ATSP:F6E21.40
 A:Experimental source: cultivar Columbia; BAC clone F6E21
 C:Genetics:
 A:Gene: ATSP:F6E21.40
 A:Map position: 4
 A:Insertions: 47/3; 87/1; 123/3; 203/3; 230/2; 255/3; 284/3; 305/1; 335/3; 347/3; 370/3; 393;
 C:Superfamily: Schizosaccharomyces pombe negative regulator of mitosis ehb1

Query Match 3.9%; Score 111; DB 2; Length 670;
 Best Local Similarity 22.4%; Pred. No. 3.7;
 Matches 123; Conservative 77; Mismatches 194; Indels 156; Gaps 33;

QY 51 LVSPSFLSTIDANLATDPRFLILGSPKRTLARGLSPAY--LRFGKTKDFLIFDPK- 107
 Db 47 LVDPSPBSLVHGN--GVDTQVLPVCGSDLV-----LSFSQSSHVVGKISSWIDDSHD 99
 QY 108 -----KESTFEERSYVQSOVNQDICKYSSIPDVEEKLRLWYOEOLLREHYOKKEX 161
 Db 100 EYLRMSDETTLKQELAMATHLSIQWCE-----PD-----LTPHYLAGGL 139
 QY 162 NSTYRSSV---DVLV-----TFPANC--GLDILFGNALLRTPADLQWSSNQML 207
 Db 140 RVSCCRSSEFISDETLKYTFNQALTFCCSSIFCLNIVSALMLWRLVPLVK--SEGDSM 196
 QY 208 LDYCSSKGYNISWEIGN-----EPNSFLKKA-DIFIN-----GSQUGEDIYQLHKL 253
 Db 197 DD--TSBGLNDSEWELMNSFRLLCEHDSKLSVALDVLTSPSTSGRWGES--VRAALIS 253
 QY 254 RKSTFNKATLYGPDPGP--RRKTKMLKSLF--KAGGEVIDSVTMHHYIANGRTATREDF 310
 Db 254 TPAFLTNAR-----GYPLSKRHQKTLAGFDHAQVVICGKPAHNLQKPLDSSSEGTG 307
 QY 311 INPDVLDIFISSVQKVFQVNVSTRPGKVMYLGETSAYGGAPLSDTFAGFMWLDKLG 370
 Db 308 KNP--LRIYLDVVAALFQMBESLSEQRIETELGYRDFLQAPLQPLMDNLEAQTYETFE-- 362
 QY 371 LSARKGIEVNRQVFFGAGNYHLDENFDPLPDYMLSLFFKKLVGTVMASVOGSKRR 429
 Db 363 ---RSVKITYQY---RAVEKALVDR---VPDEKASEL-----ITVLMVVGAGRGPLY 406
 QY 421 -ASVOGSKR--RKLRYVYLCTNTDNPRYKEGDLTYVAINLHVTK-----YLRLEY 468
 Db 407 RASLQAABETDKAKVY--AVEKNPN-----AVVTIHLVVKMGEDVVTIISCDM 455
 QY 469 PPSN--KQVYKTLRLPLGHLSSVQLNGLTLKAVDDQTLPLM--EKLPRQSSISG 523
 Db 456 RFWNAPEQADIIVSEILSGFG-----DNELSPECIDGQRPLKP--DGIS 498
 QY 524 LPAFYSFPV 533
 Db 499 IPS-STYSFT 507

RESULT 7

S32961
 hypothetical protein YBR259w - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein YBR1727

```

QY 411 KKLVTGKVLMAVSGSKRR-----KLRYLHCTNDNPRYKEDLTLYAINLHNT 461
DB 337 KRLIGKVLAVHAGLQRPRGRVARDLRIYAHCTNHNHNHYVGSITLFINHRSR 396
QY 462 KTLRLPYPSNKNQVDKYLRLPLGPHGLSKSVOLNGLTKMDDQTLPLMEKPLRPSG 521
DB 397 KKIKLAGLTKLTHVHGLTLPYQOEGLSKSVOLNGLTKMDDQTLPLMEKPLRPSG 456
QY 522 LGLPAFVSFPYIRNAKVAAC 542
DB 457 LVIPVYMGFVYKVNALAC 477

```

RESULT 2

T45608

hypothetical protein F13G24.30 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear crests)

C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004

C/Accession: T45608

R.Beaven, M.; Van Der Schueren, J.; Chuang, Y.U.; Voelt, M.; Robben, J.; Volckaert, G.; Ba submitted to the Protein Sequence Database, December 1999

A/Reference number: Z23009

A/Accession: T45608

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-521 <BEV>

A/Cross-references: UNIPROT:Q9SDA1; EMBL:AL133421

C/Genetics:

A/Map position: 5

A/Intons: 53/3; 66/1; 127/2; 177/1; 256/1; 319/2; 361/2; 394/3

A/Note: F13G24.30

Query Match 14.6%; Score 416; DB 2; Length 521;

Best Local Similarity 29.2%; Pred. No. 5.8e-23;

Matches 154; Conservative 68; Mismatches 184; Indels 122; Gaps 24;

```

QY 75 LGSPLKRLTARGLSPAYLRFSGTKTDLIFDPKKESTFEBSYWGQVNOIDICKYGSIP 134
DB 55 LTRPLTKAIAKFKPIRIGSLQDQVYIDVGNLKT-----PCR----- 94
QY 135 DVEEKRLLEMPYQEOQLLREHYOKKPKNS---TYSRSV-----DVLTYFANCSGLDLIF 186
DB 95 -----PQOKM-----NSGLGFGSKGLMKMKMDLNSFLTATGAVTF 132
QY 187 GINALILRTDLQ-----WNSNAQLLDYCSSKGTN-SMELGNPNSFLKADIFIN 238
DB 133 GINALRGRHKLKRGKAWGAMDHINTDPLNTVSKGYVDSWFGNELSG--SGVGASVS 190
QY 239 GSQLEGDIYQLHKLKSTFKNAKLYGPPVGP-----BRKTAKMLKSLKAGGEYIDSV 293
DB 191 AEYKGLIYLVKDVINK-VYKSNMHLKPLVABGGEYEQOWTKLEI---SOPSVADV 246
QY 294 TWHHYLLNGRT-ATREDFLNPVDLIDFISVQKV-----QVESSTRPGKWLGETSSA 347
DB 247 THHIYVLGSGNDPALVKIIMDP-----YLSQVSKTEKDVNOQTIGHPMASPVGSSGA 302
QY 348 YGGGALLDPTPAAGGMMLDKGLSRMGLIYVMQVFFGAGNYHLVDR-NPDPDLYYL 406
DB 303 YNSGGHNVBDTFIDSFYLDOLMSNRHNTKYVCRQTLVG-GFYGLLEKTFPNDPYS 361
QY 407 SLTFKLVTGKVLMAVSGSKRRKLVYLHCTNDNPRYKEDLTLYAINLHNTKYL-- 464
DB 362 ALLMHLKMGKVLAVTDPDP--QLRVYAHCSK-----GRACVTLLINLSQSPFTYS 413
QY 465 -----RLPYPS---NKQVDKYLIR---LGPHG--ILSKSVOL 495
DB 414 VNSGINVVLNAESRKKKSLDLTKRPFWSIGSKASGYLNRBEYHLTPENGVLRSKTMVL 473
QY 496 NGLTIKMDVDQTLPLMEKPLR-GSSGLIPAFSYPFYIRNAKVAAC 542
DB 474 NKSILKPTATGDIPL-EPVLSVNSPLNVLPLSMSTFIVLPNDASAC 520

```

RESULT 3

T01953

hypothetical protein T2L5.6 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear crests)

C/Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 09-Jul-2004

C/Accession: T01953

R.Giesel, C.; Smith, A.; Le, T.

submitted to the EMBL Data Library, October 1998

A/Description: The sequence of A. thaliana T2L5.

A/Reference number: Z14470

A/Accession: T01953

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-190 <GEI>

A/Cross-references: UNIPROT:O82604; EMBL:AF096371; NID:g3695386; PID:g3695392

A/Experimental source: cultivar Columbia

C/Genetics:

A/Map position: 4

A/Intons: 36/2; 69/3

A/Note: T2L5.6

C/Superfamily: Arabidopsis thaliana hypothetical protein T2L5.6

Query Match 6.0%; Score 169.5; DB 2; Length 190;
 Best Local Similarity 27.8%; Pred. No. 2.5e-05;
 Matches 54; Conservative 34; Mismatches 57; Indels 49; Gaps 9;

```

QY 382 ROVFGAGNYHLVD-ENFDPDLYLWLSLFFKLVGKVLMAVSGSKRRKRLVYLHCTNT 440
DB 12 RSLIG-GNYGLNTNTFTPNPDYSAIWRQLMGRKALFTTFSQTK--KIRSYTHCA-- 66
QY 441 DNPYKEDLTLYAINLHV-----TKYLRLPYPSNKNQVDKYLRL 483
DB 67 ---ROSKG-ITVLMLNDITTVAKVLANNSFSLRHKMKM-----SKRASSQLFG-- 115
QY 484 GPHGLL-----SKSVQNLGLTLKMWDDQTLPLMEKPLRPGSSGLGPAFS 528
DB 116 GNGVYIQREYHLTKADGNLSQTMNLGNALQVNSMGDLPIEPIHINSTEPIITAPYS 175
QY 529 YSFPVIRNAKVAAC 542
DB 176 IVFVHRNVVPAC 189

```

RESULT 4

F64383

hypothetical protein M0670 - Methanococcus jannaschii

C/Species: Methanococcus jannaschii

C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004

C/Accession: F64383

R.Bult, C.J.; White, O.; Olson, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, J.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weissbrock, K.G.; Merrick, J.M.; Glodek, A.; Ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A/authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.

A/Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.

A/Reference number: A64300; MUID:96337999; PMID:8688087

A/Accession: F64383

A/Status: preliminary

A/Molecule type: nucleic acid sequence not shown; translation not shown

A/Residues: 1-356 <BUL>

A/Cross-references: UNIPROT:Q58084; GB:U67514; GB:L77117; NID:g2826304; PIDV:AA89664.1;

C/Genetics:

A/Map position: REV596356-595886

A/Start codon: GTG

Query Match 4.0%; Score 112.5; DB 2; Length 356;
 Best Local Similarity 21.2%; Pred. No. 1.1;
 Matches 85; Conservative 48; Mismatches 143; Indels 125; Gaps 18;

```

QY 126 ICKY-----GSIPDVEKRLLEMPYQEOQLLREHYOKKPKNSYSSRSVD----- 171

```



GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2005, 11:19:33 ; Search time 27 seconds
(without alignments)
1935.026 Million cell updates/sec

Title: SEQ2A
Perfect score: 2842
Sequence: 1 MLRSEKPLPPLMLLLGP.....LPAFSYPFVIRNAKVAACI 543

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	897.5	31.6	480	2 JC7506	heparanase protein
2	416	14.6	521	2 T45608	hypothetical prote
3	169.5	6.0	190	2 T01953	hypothetical prote
4	112.5	4.0	356	2 F64383	hypothetical prote
5	111.5	3.9	575	2 T12094	beta-fructofuranos
6	111	3.9	670	2 T10666	hypothetical prote
7	111	3.9	688	2 S32961	hypothetical prote
8	111	3.9	2298	2 T49648	hypothetical prote
9	109.5	3.9	879	2 E91031	probable outer mem
10	108.5	3.8	411	2 S74760	hypothetical prote
11	107.5	3.8	500	2 D87541	beta-xylosidase [1
12	106	3.7	879	2 F85875	probable fibrinoly
13	105	3.7	670	2 T38446	microtubule-associ
14	104.5	3.7	788	1 S00652	phosphoribosylamin
15	104	3.7	432	2 F70411	adenylosuccinate s
16	104	3.6	2013	2 A11489	probable peptidogl
17	103.5	3.6	587	2 S36231	beta-fructofuranos
18	103.5	3.6	676	2 AFI153	transcription anti
19	103.5	3.6	687	2 F85188	retrotransposon 11
20	103	3.6	796	2 D97055	transketolase [imp
21	101	3.6	594	2 A82913	hypothetical prote
22	101	3.6	644	2 A97268	methionyl-tRNA syn
23	100.5	3.5	805	2 H72098	DNA gyrase, chain
24	100.5	3.5	805	2 C86525	DNA gyrase subunit
25	100.5	3.5	989	2 AE2140	toxin secretion AB
26	99.5	3.5	510	2 H69893	conserved hypochet
27	99.5	3.5	837	1 A31842	endo-1,4-beta-xyla
28	99	3.5	897	2 G02529	dynein heavy chain
29	99	3.5	4644	1 A38905	dynein heavy chain

30	98.5	3.5	596	2 T04506	hypothetical prote
31	98.5	3.5	629	2 C64180	hypothetical prote
32	98.5	3.5	654	2 T14202	NADH2 dehydrogenas
33	98.5	3.5	699	2 P95146	DNA topoisomerase
34	98.5	3.5	701	2 D98014	DNA topoisomerase
35	98.5	3.5	746	2 T46821	sideophore recept
36	98.5	3.5	746	2 A95420	Rhca Rhizobactin r
37	98.5	3.5	1012	2 JC5925	membrane kiocho pr
38	98	3.4	465	2 T19113	hypothetical prote
39	98	3.4	716	1 C60008	RNA-directed RNA p
40	98	3.4	760	2 T34414	hypothetical prote
41	98	3.4	817	2 H75035	probable membrane
42	97.5	3.4	454	2 T20829	probable serine ca
43	97.5	3.4	511	2 S61166	probable membrane
44	97.5	3.4	604	2 E75119	hypothetical prote
45	97.5	3.4	804	2 G71546	probable DNA gyras

ALIGNMENTS

RESULT 1

UC7506
heparanase protein 2a - human
C/Species: Homo sapiens (man)
C/Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 09-Jul-2004
C/Accession: JC7506
R/McKenzie, E., Tyson, K., Stamps, A., Smith, P., Turner, P., Barry, R., Hircock, M., Pat
Biochem. Biophys. Res. Commun. 276, 1170-1177, 2000
A/Title: Cloning and expression profiling of Hpa2, a novel mammalian heparanase family me
A/Reference number: JC7506
A/Accession: JC7506
A/Molecule type: mRNA
A/Residues: 1-480 <MKA>
A/Cross-references: UNIPROT:Q9HB39; GB:A282885
C/Comment: This protein, a intracellular membrane-bound enzyme, has biological and therat
therapies.
C/Genetics:
A/Genes: hpa2a
A/Map position: 10q23-10q24
C/Keywords: heparin binding; membrane bound

Query Match	31.6%	Score 897.5	DB 2:	Length 480;
Best Local Similarity	36.0%	Pred. No. 9.2e-59;		
Matches	202;	Conservative 74;	Mismatches 146;	Indels 139; Gaps 9
QY	20	PLGPIISPGAL-----PRPA-----QAQVVDIDFTQEPHLVSPS	55	
DB	18	PPACIAPGALYIALLLHLHLSLSQAQGRRLPVDRRAAGLKEKTLILDVSTKQNPRTVEN	77	
QY	56	PLSVTIDAMLATDPRFLILGSPKLTTLARGLSPAYLRFQGTQDFLIF----DEPKRST	111	
DB	78	FLSLQDPSIITHD-GWLDPLSSKRLVTLARGLSPALRRGCKRTDPLQONLRNPKSR-	135	
QY	112	FEERSYQSQVQNDICKYSSIPPDVEEKLRLWPOEQLLRHYQYKTRKNSYSSVSD	171	
DB	136	-----GAPGPD-----YLNKYE-----	148	
QY	172	VLTPANCGLDILFGLNALLRTRADIQWSSNAQILLDYCSSKGYNISWELGNEPNSLX	231	
DB	149	-----DEPNNYRT	156	
QY	232	KADIFNGSOLGEDIYQLHKLRLK-STFKNAXLYGPDVQOPRRKATKMLKSLKAGGEYI	290	
DB	157	MGRAYNSQLGKDIYQLKSLDPRIRIYGRASISYGNIRPRKNVYALLDGFMKAVGSTV	216	
QY	291	DSVTWRYHLYNGTATREDPLNDVDLIFISSVQKVFQVESSTRPCKKVMGLGETSSAYGG	350	
DB	217	DAVTWQHCYIDGRVAVKMPDLKTRLDLTDSDQIRKIQKVNVTYTPCKKTLWLBGVVTSAG	276	
QY	351	GALLSDTPAAGFMUDKGLSARMGIEVVMRQVFEAGANVHYVDNPPILPDYMLSLF	410	
DB	277	GTNILSDSYAAGPLMINTLIGMLANOGIDIVIRHSFPDHGNHLVDQNFPLDPMYLSLY	336	

Query Match 99.9%; Score 2838; DB 2; Length 543;
Best Local Similarity 99.8%; Pred. No. 6.2e-273;
Matches 542; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 MLRLSKPALPPPLMLLLGLPLGSPGALPRPAQADVVLDLDFTOEPHLVSPSLSVT 60
DB      1 MLRLSKPALPPPLMLLLGLPLGSPGALPRPAQADVVLDLDFTOEPHLVSPSLSVT 60
QY      61 IDANLATDRPFLILGSPKLTARGLSPAYLRFSGTKTDPLIFDPKKESTFEERSYMS 120
DB      61 IDANLATDRPFLILGSPKLTARGLSPAYLRFSGTKTDPLIFDPKKESTFEERSYMS 120
QY      121 QVNODICKYGSIPPDVEEKLRLMPYQEOQLLREHYQKKFKNSTYSSSVVLYTFPANC 180
DB      121 QVNODICKYGSIPPDVEEKLRLMPYQEOQLLREHYQKKFKNSTYSSSVVLYTFPANC 180
QY      181 GIDLTFGLNALRTADLQWSSNAOQLLDYCSSKGYNISWELGNEPNSFLKADIFINGS 240
DB      181 GIDLTFGLNALRTADLQWSSNAOQLLDYCSSKGYNISWELGNEPNSFLKADIFINGS 240
QY      241 QLGEDYIQAHKLRLKSTFGNAKLYGPDVQGPRTAKMLKSPFKAGGEVIDSVTWHYYL 300
DB      241 QLGEDYIQAHKLRLKSTFGNAKLYGPDVQGPRTAKMLKSPFKAGGEVIDSVTWHYYL 300
QY      301 NGRTATREDFLNPDVLDIFISSVQKVFQYVESTRPGKXWLGETSAYGGAPLSDTFA 360
DB      301 NGRTATREDFLNPDVLDIFISSVQKVFQYVESTRPGKXWLGETSAYGGAPLSDTFA 360
QY      361 AGFMWLDKLGASRMGI EYVMRQVFFGAGNYHVDENFPLPDYMLSLFKLVGTVKM 420
DB      361 AGFMWLDKLGASRMGI EYVMRQVFFGAGNYHVDENFPLPDYMLSLFKLVGTVKM 420
QY      421 ASVQSKRRKRLRYVLHCTNTDNPRYKEGDLTYAINLHNVTXYLRLEPYPSNKOVDKYL 480
DB      421 ASVQSKRRKRLRYVLHCTNTDNPRYKEGDLTYAINLHNVTXYLRLEPYPSNKOVDKYL 480
QY      481 RPLGPHGLLSKSVQNLGLTLKXVDDQTLPPLMKEPLRPSSSLGLPAFSYSFFVIRNAKYA 540
DB      481 RPLGPHGLLSKSVQNLGLTLKXVDDQTLPPLMKEPLRPSSSLGLPAFSYSFFVIRNAKYA 540
QY      541 ACI 543
DB      541 ACI 543
```

Search completed: March 23, 2005, 11:25:48
Job time : 117.5 secs

```

XX  US2003217375-A1.
PN
XX
XX  20-NOV-2003.
PD
XX
XX  24-FEB-2003; 2003US-00371218.
PF
XX  31-AUG-1998; 98WO-US017954.
XX  01-MAR-1999; 99US-00258892.
XX  06-FEB-2001; 2001US-00776874.
PR  19-NOV-2001; 2001US-00988113.
XX
XX  (ZCHARA/) ZCHARIA E.
PA  (VL0D/) VL0DAVSKY I.
PA  (METZ/) METZGER S.
PA  (PECK/) PECKER I.
PA  (ILAN/) ILAN N.
PA  (CHAU/) CHAUER-SHAUL T.
PA  (GOLD/) GOLDSHMIDT O.
XX
XX  Zcharia E, Vlodaysky I, Metzger S, Pecker I, Ilan N;
PI  Chajek-Shaul T, Goldshmidt O;
DR  N-PSDB; ADM48719, ADM48721.
XX
XX  WPI: 2004-021918/02.
XX
XX  New transgenic non-human animal expressing heparinase, useful as models
PT  for human disease, such as cancers, viral infection, neurodegenerative
PT  diseases, restenosis, atherosclerosis and pulmonary disorders.
XX
XX  Example 6; SEQ ID NO 14; 106pp; English.
XX
XX  The present invention relates to a transgenic non-human animal whose
CC  genome comprises an exogenous polynucleotide sequence, including a
CC  promoter active in tissues of the non-human, a region encoding a human
CC  heparanase, where the promoter and the region encoding human heparanase
CC  are operably linked in the exogenous polynucleotide such that human
CC  heparanase is expressed in at least a portion of the cells of the non-
CC  human animal. The methods and compositions of the present invention are
CC  useful for the production of transgenic animals expressing heparanase, to
CC  be used as models for human diseases such as cancers, viral infection,
CC  restenosis, neurodegenerative diseases, atherosclerosis and pulmonary
CC  disorders. The present sequence is human SK-hepi hpa protein used in the
CC  exemplification of the invention.
XX
XX  Sequence 592 AA;
SQ

```

```

Query Match      100.0%; Score 2842; DB 8; Length 592;
Best Local Similarity 100.0%; Pred. No. 2.9e-273;
Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY  1 MLRSKPAALPPPLMLLLGPIGLSPGALPRPAQADVDLDFPQEPHLVSPFLSYT 60
DB  50 MLRSKPAALPPPLMLLLGPIGLSPGALPRPAQADVDLDFPQEPHLVSPFLSYT 109
QY  61 IDANLATDRPFLILGSPKRLTARGISPAVYLFSGTKTDFLIFDPKSTPERESYWS 120
DB  110 IDANLATDRPFLILGSPKRLTARGISPAVYLFSGTKTDFLIFDPKSTPERESYWS 169
QY  121 QVNQDICKGSGIPDPVEEKLRLFWPYQEOILLREHYQKFKNSTYRSYDVLYTANCS 180
DB  170 QVNQDICKGSGIPDPVEEKLRLFWPYQEOILLREHYQKFKNSTYRSYDVLYTANCS 229
QY  181 GLDLIGLNAALFTALQWSSNAOILLDYCSSKGNISWELGNEPNSFLKKADIFINGS 240
DB  230 GLDLIGLNAALFTALQWSSNAOILLDYCSSKGNISWELGNEPNSFLKKADIFINGS 289
QY  241 QLGEDYIOLHKLKSTFKNAKLYGPDVGQPRKRTAKMLKSFKAGGEVIDSVTHHYVL 300
DB  290 QLGEDYIOLHKLKSTFKNAKLYGPDVGQPRKRTAKMLKSFKAGGEVIDSVTHHYVL 349
QY  301 NGRTATREDPLNDVLDIFISSVQKVFQVVESTRPGRKWLGETSSAYGGADPLSDTPA 360
SQ

```

```

DB  350 NGRTATREDPLNDVLDIFISSVQKVFQVVESTRPGRKWLGETSSAYGGADPLSDTPA 409
QY  361 AGFMWLDKLGSAFMGIEVYMKQVFPFAGNHYLVDENEPDLBDYMLSLFKKLVGTKVLM 420
DB  410 AGFMWLDKLGSAFMGIEVYMKQVFPFAGNHYLVDENEPDLBDYMLSLFKKLVGTKVLM 469
QY  421 ASVQGSKRRKRLRYVTHCTNTDNPRIKEGDLTYAINTLNHTYKTLALPPYFNSKQVDKYL 480
DB  470 ASVQGSKRRKRLRYVTHCTNTDNPRIKEGDLTYAINTLNHTYKTLALPPYFNSKQVDKYL 529
QY  481 RPLGPHGLSKSVOLNGTLTKWVDOTLPPLMEKPLRPSSSLGLPAFSYFVIRNAKYA 540
DB  530 RPLGPHGLSKSVOLNGTLTKWVDOTLPPLMEKPLRPSSSLGLPAFSYFVIRNAKYA 589
QY  541 ACT 543
DB  590 ACT 592

```

RESULT 15
AA17082
ID AA17082 standard; protein; 543 AA.
XX
XX AA17082;
XX
XX 21-JUL-1999 (first entry)
XX
DE Human heparanase enzyme.
XX
XX Heparanase; endoglucuronidase; heparan sulfate proteoglycan; enzyme;
XX metastasis; angiogenesis; wound healing; angioplasia-induced restenosis;
XX arteriosclerosis; atherosclerosis; inflammation; tissue development;
XX human; HSPG.
XX
XX Homo sapiens.
XX
XX WO921975-A1.
XX
XX 06-MAY-1999.
XX
XX 28-OCT-1998; 98WO-AU000898.
XX
XX 28-OCT-1997; 97AU-00000062.
XX 09-DEC-1997; 97AU-00000812.
XX
XX (AUSU) UNIV AUSTRALIAN NAT.
XX
XX Freeman CG, Hulett MD, Parish CR, Hamdorf BJ;
PI WPI: 1999-312956/26.
XX N-PSDB; AAX37259.
DR
XX Polynucleotides encoding mammalian endoglucuronidases, especially
PT heparanases, useful to promote wound healing.
XX
XX Claim 6; Page 69-73; 112pp; English.
XX
XX The invention relates to nucleic acid sequences that encode heparanase
CC enzymes having endoglucuronidase activity. Recombinant heparanases are
CC capable of removing the HS side chain from heparan sulfate proteoglycan
CC (HSPG). Sulfated oligosaccharides, sulphates or HSPG can be used to
CC inhibit heparanase, this is useful for treatment of a physiological or
CC medical condition associated with elevated heparanase activity, such as
CC metastasis, angiogenesis, wound healing, angioplasia-induced restenosis,
CC arteriosclerosis, atherosclerosis and inflammation. The human, murine and
CC rat heparanases can be used to enhance wound healing, especially
CC associated with tissue development and repair. The conditions mentioned
CC above can be diagnosed using specific antibodies, and also using primers
CC and probes specific for the heparanase polynucleotides. Other uses of the
CC heparanases include sequencing sulfated molecules such as HSPG. The
CC present sequence represents a human heparanase
SQ Sequence 543 AA;


```

QY 361 AGFMWLDKLGISARNGIEVVMRQVFFGAGNYHLVDENFDPLPDYWSLLFKKLVGTKVLM 420
DB 410 AGFMWLDKLGISARNGIEVVMRQVFFGAGNYHLVDENFDPLPDYWSLLFKKLVGTKVLM 469
QY 421 ASVQGSKRKRKLAVYLHCTNTDNPARYKSGDLTLTYAINLHVNTKYLRLPYFSKNQVDKYL 480
DB 470 ASVQGSKRKRKLAVYLHCTNTDNPARYKSGDLTLTYAINLHVNTKYLRLPYFSKNQVDKYL 529
QY 481 RPLGPHGLSKSVQNLGLTLKMWDDOTLPLMEKPLRPSSISGLPAFYSFVIRNAKVA 540
DB 530 RPLGPHGLSKSVQNLGLTLKMWDDOTLPLMEKPLRPSSISGLPAFYSFVIRNAKVA 589
QY 541 ACT 543
DB 590 ACT 592

RESULT 13
ADL16383
ID ADL16383 standard; protein; 592 AA.
XX
AC ADL16383;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human heparanase full length protein.
XX
KM Human; heparanase; heparanase-dependent cancer; cancer;
XX autoimmune reaction; inflammation; chromosome 4; enzyme.
XX
OS Homo sapiens.
XX
PN US200326215-A1.
XX
PD 25-DEC-2003.
XX
PF 09-JUN-2003; 2003US-00456573.
XX
PR 31-AUG-1998; 98WO-US017954.
PR 01-MAR-1999; 99US-00256892.
PR 08-NOV-1999; 99US-00435739.
XX
PA (INSI-) INSIGHT STRATEGY & MARKETING LTD.
PA (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
XX
PI Pecker I, Vlodaevsky I, Feinstein E;
XX
DR WPI; 2004-070610/07.
XX
PT New antisense oligonucleotide hybridizable with a polynucleotide encoding
PT a polypeptide with heparanase activity, useful for treating diseases such
PT as cancer and autoimmune disorders.
XX
PS Claim 3; SEQ ID NO 14; 108bp; English.
XX
CC The invention relates to an antisense oligonucleotide (ASO) comprising a
CC polynucleotide or a polynucleotide analogue of at least 10 bases being
CC hybridizable in vivo, under physiological conditions, with a portion of
CC a polynucleotide strand encoding a polypeptide having heparanase
CC catalytic activity. Also included are a method of in vivo downregulating
CC heparanase activity (comprising administering the ASO in vivo), a method
CC of treating a subject suffering from a pathological condition
CC (characterised by heparanase activity, comprising administering ASO to
CC the subject), a pharmaceutical composition comprising the ASO and a
CC carrier, an antisense nucleic acid construct (comprising a promoter
CC sequence and a polynucleotide sequence directing the synthesis of an
CC antisense RNA sequence of at least 10 bases being hybridizable in vivo,
CC under physiological conditions, with a polynucleotide strand encoding a
CC polypeptide having heparanase catalytic activity), a method of in vivo
CC downregulating heparanase activity (comprising administering in vivo the
CC antisense nucleic acid construct), a pharmaceutical composition
CC comprising the antisense nucleic acid construct and a carrier, and an
CC antisense oligonucleotide comprising a polynucleotide or a polynucleotide

```

```

CC analogue of at least 10 bases being hybridizable in vivo, under
CC physiological conditions, with a portion of a polynucleotide strand being
CC characterised by forming at least a portion of an untranslated region
CC (UTR) for a polynucleotide strand encoding a polypeptide having
CC heparanase catalytic activity. The methods and compositions of the
CC present invention are useful for the prevention and/or treatment of
CC diseases or conditions associated with aberrant heparanase activity, such
CC as heparanase-dependent cancer, cancer, autoimmune reaction and
CC inflammation. The gene for human heparanase is located on chromosome 4.
CC The present sequence is a human heparanase protein.
XX
SQ Sequence 592 AA;
XX
Query Match 100.0%; Score 2842; DB 8; Length 592;
Best Local Similarity 100.0%; Pred. No. 2.9e-273;
Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLRSKPALPPPLMLLLGLGSPGLPRPAQADVVDDFFTOEPLHLVSPSFLSYT 60
DB 50 MLRSKPALPPPLMLLLGLGSPGLPRPAQADVVDDFFTOEPLHLVSPSFLSYT 109
QY 61 IDANLATDPRFLILGSPKLTARGLSPAYLRFSGTKTDPLIPDKKSTFEERSYWS 120
DB 110 IDANLATDPRFLILGSPKLTARGLSPAYLRFSGTKTDPLIPDKKSTFEERSYWS 169
QY 121 QVNODICKYGSIPDVEEKLRLMPYOBQLLREHYOKKFNKSTYSRSSVDLYTFPANC 180
DB 170 QVNODICKYGSIPDVEEKLRLMPYOBQLLREHYOKKFNKSTYSRSSVDLYTFPANC 229
QY 181 GIDLIFGNALIRTDLDQNSSNAQLLDYCSSKGNYSWEIGNEPNFLKKADIIFINCS 240
DB 230 GIDLIFGNALIRTDLDQNSSNAQLLDYCSSKGNYSWEIGNEPNFLKKADIIFINCS 289
QY 241 QLGEDYIOHLKLRSTFNNAKLYGPDVGQPRKTAAMKSLFKXGGEYIDVTHHHYL 300
DB 290 QLGEDYIOHLKLRSTFNNAKLYGPDVGQPRKTAAMKSLFKXGGEYIDVTHHHYL 349
QY 301 NGRTATREDPLNPDVLDIFISSVQKVFQVVESTRGKKVWLGETSSAYGGAPLLSDTFA 360
DB 350 NGRTATREDPLNPDVLDIFISSVQKVFQVVESTRGKKVWLGETSSAYGGAPLLSDTFA 409
QY 361 AGFMWLDKLGISARNGIEVVMRQVFFGAGNYHLVDENFDPLPDYWSLLFKKLVGTKVLM 420
DB 410 AGFMWLDKLGISARNGIEVVMRQVFFGAGNYHLVDENFDPLPDYWSLLFKKLVGTKVLM 469
QY 421 ASVQGSKRKRKLAVYLHCTNTDNPARYKSGDLTLTYAINLHVNTKYLRLPYFSKNQVDKYL 480
DB 470 ASVQGSKRKRKLAVYLHCTNTDNPARYKSGDLTLTYAINLHVNTKYLRLPYFSKNQVDKYL 529
QY 481 RPLGPHGLSKSVQNLGLTLKMWDDOTLPLMEKPLRPSSISGLPAFYSFVIRNAKVA 540
DB 530 RPLGPHGLSKSVQNLGLTLKMWDDOTLPLMEKPLRPSSISGLPAFYSFVIRNAKVA 589
QY 541 ACT 543
DB 590 ACT 592

RESULT 14
ADMA8720
ID ADMA8720 standard; protein; 592 AA.
XX
AC ADMA8720;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human SK-hep1 hpa protein.
XX
KM Transgenic animal; heparanase; cancer; viral infection; restenosis;
XX neurodegenerative disease; atherosclerosis; pulmonary disorder; hpa;
XX human.
XX
OS Homo sapiens.

```

XX Claim 22; Page 122-123; 152pp; English.
PS The present sequence represents a human protein with heparanase catalytic
XX activity. The heparanase (hpa) polynucleotide is useful in gene therapy.
CC Particularly in treating tumour, inflammation or autoimmunity.
CC Particularly, the polynucleotide is useful in modulating the
CC bioavailability of heparin-binding growth factors, cellular responses to
CC heparin-binding growth factors (e.g. bFGF) and cytokines (e.g.
CC interleukin (IL)-6), cell interaction with plasma lipoproteins, cellular
CC susceptibility to certain viral and some bacterial and protozoa
CC infections, or disintegration of neurodegenerative plaques. The
CC polynucleotide is also useful in wound healing (e.g. thermal, chemical or
CC radiation burns), and in the treatment of angiogenesis, restenosis,
CC atherosclerosis, inflammation, neurodegenerative diseases (Gerstmann-
CC Strausler Syndrome or Creutzfeldt-Jakob disease), and some viral,
CC bacterial or protozoa infections
SQ Sequence 592 AA;
Query Match 100.0%; Score 2842; DB 3; Length 592;
Best Local Similarity 100.0%; Pred. No. 2.9e-273; Indels 0; Gaps 0;
Matches 543; Conservative 0; Mismatches 0;
QY 1 MLRSKPALPPMLMLLGLPLSPGALPRPAQADVVDLDFTEPHLVSPFLSVT 60
DB MLRSKPALPPMLMLLGLPLSPGALPRPAQADVVDLDFTEPHLVSPFLSVT 109
QY 61 IDANLATDRPFLILGSPKRLTLARGLSPAYLRFGGTKTDPLIFDPKKSSTFEERSYWS 120
DB IDANLATDRPFLILGSPKRLTLARGLSPAYLRFGGTKTDPLIFDPKKSSTFEERSYWS 169
QY 110 IDANLATDRPFLILGSPKRLTLARGLSPAYLRFGGTKTDPLIFDPKKSSTFEERSYWS 169
DB 121 QVNMDCIKGSIPTDVEEKLRLMPYQEOULLREHQQKFKNSTYRSADVLYTPANC 180
QY 170 QVNMDCIKGSIPTDVEEKLRLMPYQEOULLREHQQKFKNSTYRSADVLYTPANC 229
DB 181 GLDLIFGLNALRTADLQWNSNAQLLDYCSKGYNISWELGNEBNSFLKADIFINGS 240
QY 230 GLDLIFGLNALRTADLQWNSNAQLLDYCSKGYNISWELGNEBNSFLKADIFINGS 289
DB 241 QUGEDYIQLHKLIRKSTFKNAKLYGPDVQPRKRTAKMLKSLKAGGEVIDSTVWHYYL 300
QY 290 QUGEDYIQLHKLIRKSTFKNAKLYGPDVQPRKRTAKMLKSLKAGGEVIDSTVWHYYL 349
DB 301 NGRTATREDPLNDVDLDFISSVQKVFQVVESTRPGKRWLGETSSAYGGAPLSDTPA 360
QY 350 NGRTATREDPLNDVDLDFISSVQKVFQVVESTRPGKRWLGETSSAYGGAPLSDTPA 409
DB 361 AGFMWLDKLGSLARMGIEVVRQVFFGAGNYHLVDENFDPLPDYMLSLLEFKLVGTVLM 420
QY 410 AGFMWLDKLGSLARMGIEVVRQVFFGAGNYHLVDENFDPLPDYMLSLLEFKLVGTVLM 469
DB 421 ASVQSGRRRLRYLHCTNDNPRYKEGDLTYAINLHNTKTLRLPYPSNQVDKYL 480
QY 470 ASVQSGRRRLRYLHCTNDNPRYKEGDLTYAINLHNTKTLRLPYPSNQVDKYL 529
DB 481 RPLGPHGLSKSVQNLGLTLKWVDDOTLPLMEKPLRPGSSGLPAFYSFFVIRNAKVA 540
QY 530 RPLGPHGLSKSVQNLGLTLKWVDDOTLPLMEKPLRPGSSGLPAFYSFFVIRNAKVA 589
DB 541 ACI 543
QY ACI
DB 590 ACI 592

RESULT 12
ADG88804
ID ADG88804 standard; protein; 592 AA.

XX ADG88804;
XX
XX 11-MAR-2004 (first entry)
XX

DE Human SK-hep1 protein.
XX Wound healing; heparanase; ulcer; burn; laceration; surgical incision;
KW necrosis; pressure wound; diabetic ulcer; angiogenesis; human; therapy.
XX Homo sapiens.
XX US2003161823-A1.
XX 28-AUG-2003.
XX 14-JAN-2003; 2003US-00341582.
XX 31-AUG-1998; 98WO-US017954.
PR 01-MAR-1999; 99US-00258892.
PR 06-FEB-2001; 2001US-00776874.
PR 05-SEP-2001; 2001WO-1L000830.
PR 19-NOV-2001; 2001US-00988113.
XX (ILAN/) ILAN N.
PA (VLAD/) VLADAVSKY I.
PA (YACO/) YACOBY-ZEEVI O.
PA (PECK/) PECKER I.
PA (FEIN/) FEINSTEIN E.
XX Ian N, Vladavsky I, Yacoby-Zeevi O, Pecker I, Feinstein E;
PI WPI; 2003-897910/82.
DR N-PSDB; ADG88803, ADG88805.
XX Composition for treating a wound comprising recombinant heparanase is
PT useful to induce or accelerate wound healing and induce or accelerate
PT angiogenesis.
XX Claim 2; SEQ ID NO 14; 143pp; English.
PS The present invention relates to methods and compositions for inducing
XX and/or accelerating wound healing via the catalytic activity of
XX heparanase. The invention is used to induce or accelerate a healing
CC process, particularly of an ulcer, burn, laceration, surgical incision,
CC necrosis, pressure wound, diabetic ulcer and to induce or accelerate
CC angiogenesis. The present sequence is human SK-hep1 protein.
XX
SQ Sequence 592 AA;
Query Match 100.0%; Score 2842; DB 7; Length 592;
Best Local Similarity 100.0%; Pred. No. 2.9e-273; Indels 0; Gaps 0;
Matches 543; Conservative 0; Mismatches 0;
QY 1 MLRSKPALPPMLMLLGLPLSPGALPRPAQADVVDLDFTEPHLVSPFLSVT 60
DB MLRSKPALPPMLMLLGLPLSPGALPRPAQADVVDLDFTEPHLVSPFLSVT 109
QY 61 IDANLATDRPFLILGSPKRLTLARGLSPAYLRFGGTKTDPLIFDPKKSSTFEERSYWS 120
DB IDANLATDRPFLILGSPKRLTLARGLSPAYLRFGGTKTDPLIFDPKKSSTFEERSYWS 169
QY 110 IDANLATDRPFLILGSPKRLTLARGLSPAYLRFGGTKTDPLIFDPKKSSTFEERSYWS 169
DB 121 QVNMDCIKGSIPTDVEEKLRLMPYQEOULLREHQQKFKNSTYRSADVLYTPANC 180
QY 170 QVNMDCIKGSIPTDVEEKLRLMPYQEOULLREHQQKFKNSTYRSADVLYTPANC 229
DB 181 GLDLIFGLNALRTADLQWNSNAQLLDYCSKGYNISWELGNEBNSFLKADIFINGS 240
QY 230 GLDLIFGLNALRTADLQWNSNAQLLDYCSKGYNISWELGNEBNSFLKADIFINGS 289
DB 241 QUGEDYIQLHKLIRKSTFKNAKLYGPDVQPRKRTAKMLKSLKAGGEVIDSTVWHYYL 300
QY 290 QUGEDYIQLHKLIRKSTFKNAKLYGPDVQPRKRTAKMLKSLKAGGEVIDSTVWHYYL 349
DB 301 NGRTATREDPLNDVDLDFISSVQKVFQVVESTRPGKRWLGETSSAYGGAPLSDTPA 360
QY 350 NGRTATREDPLNDVDLDFISSVQKVFQVVESTRPGKRWLGETSSAYGGAPLSDTPA 409
DB

QY	541	ACT	543
Db	541	ACT	543

RESULT 10
AAV02346
ID AAV02346 standard; protein; 592 AA.

AC	AAV02346;	
XX		
DT	09-JUL-1999	(first entry)
XX		
DE	A human heparanase protein.	

KM Heparinase; hpa; modulator; heparin-binding growth factor;
KM cellular response; cytokine; cell interaction; plasma lipoprotein;
KM cellular susceptibility; infection; disintegration;
KM neurodegenerative plaque; wound healing; angiogenesis; restenosis;
KM atherosclerosis; inflammation; neurodegenerative disease; neuritis;
KM plasma heparin; micrometastasis; autoimmune lesion; renal failure;

OS	Homo sapiens.
XX	
PN	W09911798-A1.

PD 11-MAR-1999.

PF 31-AUG-1998; 98WO-US017954.

PR	02-SEP-1997;	97US-00922170.
PR	02-JUL-1998;	98US-00109386.

PA (INST-) INSIGHT STRATEGY & MARKETING LTD.
PA (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
PA (FRIE/) FRIEDMAN M M.

PI Pecker I, Vlodavsky I, Feinstein E;

DR WPI; 1999-302255/25.
DR N-PSDB; AAX35650.

PT New human polynucleotide useful for treating angiogenesis, restenosis, PT and inflammation.

PS Claim 6; Page 65-66; 63pp; English.

CC The specification describes a polypeptide having heparanase (hpa)
CC activity. The recombinant protein is used as a modulator of heparin-
CC binding growth factors, cellular responses to heparin-binding growth
CC factors and cytokines, cell interaction with plasma lipoproteins,
CC cellular susceptibility to viral, protozoal and bacterial infections or
CC disintegration of neurodegenerative plaques. Heparanase may be useful for
CC conditions such as wound healing, angiogenesis, restenosis,
CC atherosclerosis, inflammation, neurodegenerative diseases, and viral
CC infections. Mammalian heparanase can be used to neutralize plasma
CC heparin, and anti-heparanase antibodies may be applied for
CC immunodetection and diagnosis of micrometastases, autoimmune lesions, and
CC renal failure in biopsy specimens, plasma samples, and body fluids. The
CC present sequence represents human heparanase

SQ Sequence 592 AA;

Query Match:	100.0%;	Score 2842;	DB 2;	Length 592;
Best Local Similarity	100.0%;	Pred. No. 2.9e-273;		
Matches 543; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 MLRLSRPALPPMLMLLLGLGLPSLGGALPRPAQADVDDLDFFQTQEPHLVSPSLSYT 60

Db 50 MLRLSRPALPPMLMLLLGLGLPSLGGALPRPAQADVDDLDFFQTQEPHLVSPSLSYT 109

QY 61 IDANLATDRFLILGLSPKURTIARGLSPAYLRFPGTKTDFLLIPDKESSTFEERSWQS 120

Db	110	IDANLMTDPRFLILLLGSRKRLTLAGLSBPAYLRPGGTITDPLIFDPKXKSTFEERSRWMS	169
Qy	121	QVNODICKYSGSIPDVEEKLRLMEWPYOELLIREHYOKKEKNSTYRSRSSVDVLYTFPANC	180
Db	170	QVNODICKYSGSIPDVEEKLRLMEWPYOELLIREHYOKKEKNSTYRSRSSVDVLYTFPANC	229
Qy	181	GLDLIFGLNALRLRTADLQWNSNNAQLLLDYCSSKCGNISWELGNEBNSFLKKADIFING	240
Db	230	GUDLIFGLNALRLRTADLQWNSNNAQLLLDYCSSKCGNISWELGNEBNSFLKKADIFING	299
Qy	241	QLGEBDYIQLHKLRLKRSFTFNAKLYGPDVGOFRKTKAKMLKSEFLKAGGEYIDSVTWHYYL	300
Db	290	QLGEBDYIQLHKLRLKRSFTFNAKLYGPDVGOFRKTKAKMLKSEFLKAGGEYIDSVTWHYYL	349
Qy	301	NGRTATREDFLNPBDVLDIFISSQCVFQVVESTRGCKYWLGEFTSSAYGGAPLLSDTFA	360
Db	350	NGRTATREDFLNPBDVLDIFISSQCVFQVVESTRGCKYWLGEFTSSAYGGAPLLSDTFA	409
Qy	361	AGFWMLDKLGLSARMGIEVYVMQVFEFGAGNYHVDENFDPLEPDYWLSSLFKKLVTGKVLW	420
Db	410	AGFWMLDKLGLSARMGIEVYVMQVFEFGAGNYHVDENFDPLEPDYWLSSLFKKLVTGKVLW	469
Qy	421	ASVQSSKRRKRLRVYLACTNTDNPRLKEGDLTYALNLHNTYKTLRLPYFPSSNQVDKLYL	480
Db	470	ASVQSSKRRKRLRVYLACTNTDNPRLKEGDLTYALNLHNTYKTLRLPYFPSSNQVDKLYL	529
Qy	481	RPLGHHGLSKSVQNLGLTKLVNDQUTLPLMEKXLRPPSSIGLPAFSYSPFVIRAKYA	540
Db	530	RPLGHHGLSKSVQNLGLTKLVNDQUTLPLMEKXLRPPSSIGLPAFSYSPFVIRAKYA	589
Qy	541	ACTI 543	
Db	590	ACTI 592	

RESULT 11
AAB08850
ID AAB08850 standard; protein; 592 AA

AC AAB08850;

DT 15-JAN-2001 (first entry)

DE Amino acid sequence of a human heparanase polypeptide.

KM Human, heparanase; gene therapy; tumour inflammation; autoimmunity;
KM heparin-binding growth factor; cytokine, neurodegenerative plaque;
KM wound healing; infection; burn; angiogenesis; restenosis;
KM atherosclerosis; inflammation; neurodegenerative disease;
KM Gerstmann-Sträussler Syndrome; Creutzfeldt-Jakob disease.

OS Homo sapiens.

PN W0200052178-A1.

PD 08-SEP-2000.

PF 14-FEB-2000; 2000WO-US003542.

PR 01-MAR-1999; 99US-00258892.

PA	(INSI-)	INSIGHT STRATEGY & MARKETING LTD.
PA	(HADA-)	HADASIT MEDICAL RES SERVICES & DEV
PA	(FRIE/)	FRIEDMAN M M.

Pecker I, Vlodavsky I, Feinstein E;

DR WPI; 2000-579289/54.
DR N-PSDB; AAA75053.

PT New polynucleotides encoding a polypeptide having heparanase activity
PT useful in wound healing and in gene therapy, particularly in treating
PT tumor, inflammation, autoimmune, neurodegenerative diseases.

Best Local Similarity 100.0%; Pred. No. 2.5e-273;
Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MLRSKPALPPMLMLLLGLPLSPGALPRPAQADVDVLDLFTQEPHLHVSPEFLSVT 60
DB 1 MLRSKPALPPMLMLLLGLPLSPGALPRPAQADVDVLDLFTQEPHLHVSPEFLSVT 60
QY 61 IDANLATDPRFLILGSPKLTARLGLSPAYLRFGGKTDFLIPDKKESTEESSYMS 120
DB 61 IDANLATDPRFLILGSPKLTARLGLSPAYLRFGGKTDFLIPDKKESTEESSYMS 120
QY 121 QVNODICKYGSIPDVEEKLRLWPYQEQLLREHYQKKFKNSTYSSRSSVDVLYTFPANC 180
DB 121 QVNODICKYGSIPDVEEKLRLWPYQEQLLREHYQKKFKNSTYSSRSSVDVLYTFPANC 180
QY 181 GIDLIFGLNALLRITADLQWSSNAQLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
DB 181 GIDLIFGLNALLRITADLQWSSNAQLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
QY 241 QUGEDYIQLHKLIRKSTFKNALYGPVVGQPPRKTKAKMLKSLKAGGEVIDSVTHHYYL 300
DB 241 QUGEDYIQLHKLIRKSTFKNALYGPVVGQPPRKTKAKMLKSLKAGGEVIDSVTHHYYL 300
QY 301 NGRTATREDPLNDVLDIFISSYQKVFQVVESTRPCKKVMLGETSAYGGAPLISDTPA 360
DB 301 NGRTATREDPLNDVLDIFISSYQKVFQVVESTRPCKKVMLGETSAYGGAPLISDTPA 360
QY 361 AGFMWLDKGLSARWGIEVVMRQVFFGAQNYHLVDENPDLDPYWLISLLFKKLVGTVKVL 420
DB 361 AGFMWLDKGLSARWGIEVVMRQVFFGAQNYHLVDENPDLDPYWLISLLFKKLVGTVKVL 420
QY 421 ASVQSKRRRLRYLHCTNTDNPRYKESGLTYALNLHVTXYLRPYPSNKQVDKYL 480
DB 421 ASVQSKRRRLRYLHCTNTDNPRYKESGLTYALNLHVTXYLRPYPSNKQVDKYL 480
QY 481 RPLGPHGLSKSVQNLGLTLKAVDDOTLPLMEKPLRPSSSLGLPAFSYSFVIRNAKYA 540
DB 481 RPLGPHGLSKSVQNLGLTLKAVDDOTLPLMEKPLRPSSSLGLPAFSYSFVIRNAKYA 540
QY 541 ACI 543
DB 541 ACI 543

RESULT 9
ADM48716
ID ADM48716 standard; protein; 543 AA.
XX
AC ADM48716;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human hpa protein #1.
XX
KW Transgenic animal; heparanase; cancer; viral infection; restenosis;
KW neurodegenerative disease; atherosclerosis; pulmonary disorder; hpa;
KW human.
OS Homo sapiens.
XX
PN US2003217375-A1.
XX
PD 20-NOV-2003.
XX
PF 24-FEB-2003; 2003US-00371218.
XX
PR 31-AUG-1998; 98WO-US017954.
PR 01-MAR-1999; 99US-00258892.
PR 06-FEB-2001; 2001US-00776874.
PR 19-NOV-2001; 2001US-00988113.
XX
XX (ZCHA/) ZCHARIA E.
PA (VL0D/) VL0DAVSKY I.
```

PA (METZ/) METZGER S.
PA (PECK/) PECKER I.
PA (ILAN/) ILAN N.
PA (CHAJ/) CHAJEK-SHAUL T.
PA (GOLD/) GOLDSMIDT O.
PI Zcharia E, Vlodavsky I, Metzger S, Pecker I, Ilan N,
PI Chajek-Shaul T, Goldsmidt O;
DR MPI: 2004-021918/02.
DR N-PEDB; ADM48715, ADM48717.

PT New transgenic non-human animal expressing heparinase, useful as models
PT for human disease, such as cancer, viral infection, neurodegenerative
PT diseases, restenosis, atherosclerosis and pulmonary disorders.

PS Example 1; SEQ ID NO 10; 106pp; English.

CC The present invention relates to a transgenic non-human animal whose
CC genome comprises an exogenous polynucleotide sequence, including a
CC promoter active in tissues of the non-human, a region encoding a human
CC heparanase, where the promoter and the region encoding human heparanase
CC are operably linked in the exogenous polynucleotide such that human
CC heparanase is expressed in at least a portion of the cells of the non-
CC human animal. The methods and compositions of the present invention are
CC useful for the production of transgenic animals expressing heparanase, to
CC be used as models for human diseases such as cancer, viral infection,
CC restenosis, neurodegenerative diseases, atherosclerosis and pulmonary
CC disorders. The present sequence is human hpa protein used in the
CC exemplification of the invention.

SQ Sequence 543 AA;

Query Match 100.0%; Score 2842; DB 8; Length 543;
Best Local Similarity 100.0%; Pred. No. 2.5e-273;
Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MLRSKPALPPMLMLLLGLPLSPGALPRPAQADVDVLDLFTQEPHLHVSPEFLSVT 60
DB 1 MLRSKPALPPMLMLLLGLPLSPGALPRPAQADVDVLDLFTQEPHLHVSPEFLSVT 60
QY 61 IDANLATDPRFLILGSPKLTARLGLSPAYLRFGGKTDFLIPDKKESTEESSYMS 120
DB 61 IDANLATDPRFLILGSPKLTARLGLSPAYLRFGGKTDFLIPDKKESTEESSYMS 120
QY 121 QVNODICKYGSIPDVEEKLRLWPYQEQLLREHYQKKFKNSTYSSRSSVDVLYTFPANC 180
DB 121 QVNODICKYGSIPDVEEKLRLWPYQEQLLREHYQKKFKNSTYSSRSSVDVLYTFPANC 180
QY 181 GIDLIFGLNALLRITADLQWSSNAQLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
DB 181 GIDLIFGLNALLRITADLQWSSNAQLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
QY 241 QUGEDYIQLHKLIRKSTFKNALYGPVVGQPPRKTKAKMLKSLKAGGEVIDSVTHHYYL 300
DB 241 QUGEDYIQLHKLIRKSTFKNALYGPVVGQPPRKTKAKMLKSLKAGGEVIDSVTHHYYL 300
QY 301 NGRTATREDPLNDVLDIFISSYQKVFQVVESTRPCKKVMLGETSAYGGAPLISDTPA 360
DB 301 NGRTATREDPLNDVLDIFISSYQKVFQVVESTRPCKKVMLGETSAYGGAPLISDTPA 360
QY 361 AGFMWLDKGLSARWGIEVVMRQVFFGAQNYHLVDENPDLDPYWLISLLFKKLVGTVKVL 420
DB 361 AGFMWLDKGLSARWGIEVVMRQVFFGAQNYHLVDENPDLDPYWLISLLFKKLVGTVKVL 420
QY 421 ASVQSKRRRLRYLHCTNTDNPRYKESGLTYALNLHVTXYLRPYPSNKQVDKYL 480
DB 421 ASVQSKRRRLRYLHCTNTDNPRYKESGLTYALNLHVTXYLRPYPSNKQVDKYL 480
QY 481 RPLGPHGLSKSVQNLGLTLKAVDDOTLPLMEKPLRPSSSLGLPAFSYSFVIRNAKYA 540
DB 481 RPLGPHGLSKSVQNLGLTLKAVDDOTLPLMEKPLRPSSSLGLPAFSYSFVIRNAKYA 540
```

```

PR 01-MAR-1999; 99US-00258892.
PR 06-FEB-2001; 2001US-0076874.
PR 05-SEP-2001; 2001WO-11000830.
PR 19-NOV-2001; 2001US-00988113.
XX
XX (ILAN/) ILAN N.
XX PA (VLAD/) VLADAVSKY I.
XX PA (YACOB/) YACOB-ZEEVI O.
XX PA (PECK/) PECKER I.
XX PA (FEIN/) FEINSTEIN E.
XX
XX Ilan N, Vladavsky I, Yacoby-Zeevi O, Pecker I, Feinstein E;
XX
XX DR MPI: 2003-897910/82.
XX N-PSDB: ADG88799, ADG88801, ADG88832.
XX
XX Composition for treating a wound comprising recombinant heparanase is
XX useful to induce or accelerate wound healing and induce or accelerate
XX angiogenesis.
XX
XX Claim 2: SEQ ID NO 10; 143bp; English.
XX
XX The present invention relates to methods and compositions for inducing
XX and/or accelerating wound healing via the catalytic activity of
XX heparanase. The invention is used to induce or accelerate a healing
XX process, particularly of an ulcer, burn, laceration, surgical incision,
XX necrosis, pressure wound, diabetic ulcer and to induce or accelerate
XX angiogenesis. The present sequence is human hpa protein.
XX
XX SQ Sequence 543 AA;
XX
XX Query Match 100.0%; Score 2842; DB 7; Length 543;
XX Best Local Similarity 100.0%; Pred. No. 2, 5e-273;
XX Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 MLRSKPALPPMLLLPLGLSPGALPRPAQADVVLDLFTQEPHLVSPFLSVT 60
XX 1 MLRSKPALPPMLLLPLGLSPGALPRPAQADVVLDLFTQEPHLVSPFLSVT 60
XX
XX 61 IDANLATDRPFLILGSPKLTARGLSPAYLRFSGTKTDFLIPDKKSTPEERSYMS 120
XX 61 IDANLATDRPFLILGSPKLTARGLSPAYLRFSGTKTDFLIPDKKSTPEERSYMS 120
XX
XX 121 QVNODICKGSIIPDVEEKLRLEMPYQEOILLREHQQKFKNSTYRSSVDVLYTANCS 180
XX 121 QVNODICKGSIIPDVEEKLRLEMPYQEOILLREHQQKFKNSTYRSSVDVLYTANCS 180
XX
XX 121 QVNODICKGSIIPDVEEKLRLEMPYQEOILLREHQQKFKNSTYRSSVDVLYTANCS 180
XX 121 QVNODICKGSIIPDVEEKLRLEMPYQEOILLREHQQKFKNSTYRSSVDVLYTANCS 180
XX
XX 181 GDLIFGLNALRTADLQWNSSNAQLLDYCSGKYNISWELGNEPNSFLKADIFINGS 240
XX 181 GDLIFGLNALRTADLQWNSSNAQLLDYCSGKYNISWELGNEPNSFLKADIFINGS 240
XX
XX 241 QUGEDVIOQHKLKRSKTFKNAKLYGPDVQPRKTKMKLSFKAGGEVIDSTMHNYVL 300
XX 241 QUGEDVIOQHKLKRSKTFKNAKLYGPDVQPRKTKMKLSFKAGGEVIDSTMHNYVL 300
XX
XX 241 QUGEDVIOQHKLKRSKTFKNAKLYGPDVQPRKTKMKLSFKAGGEVIDSTMHNYVL 300
XX 241 QUGEDVIOQHKLKRSKTFKNAKLYGPDVQPRKTKMKLSFKAGGEVIDSTMHNYVL 300
XX
XX 301 NGRTATREDPLNDVDLFISSVQKVFQVVESTRPKKVMLGHTSAGGAPLSDTPA 360
XX 301 NGRTATREDPLNDVDLFISSVQKVFQVVESTRPKKVMLGHTSAGGAPLSDTPA 360
XX
XX 301 NGRTATREDPLNDVDLFISSVQKVFQVVESTRPKKVMLGHTSAGGAPLSDTPA 360
XX 301 NGRTATREDPLNDVDLFISSVQKVFQVVESTRPKKVMLGHTSAGGAPLSDTPA 360
XX
XX 361 AGFMWLDKLGSLARMGIEVVMQVFFGAGNYHLVDENFPLPDYWLSLFKKLVTGKVL 420
XX 361 AGFMWLDKLGSLARMGIEVVMQVFFGAGNYHLVDENFPLPDYWLSLFKKLVTGKVL 420
XX
XX 361 AGFMWLDKLGSLARMGIEVVMQVFFGAGNYHLVDENFPLPDYWLSLFKKLVTGKVL 420
XX 361 AGFMWLDKLGSLARMGIEVVMQVFFGAGNYHLVDENFPLPDYWLSLFKKLVTGKVL 420
XX
XX 421 ASVQSKRRKRLRYLHCTNTDNRKYEGDLTYAIVLHNVTKLRLPYPSNKNQVVKYL 480
XX 421 ASVQSKRRKRLRYLHCTNTDNRKYEGDLTYAIVLHNVTKLRLPYPSNKNQVVKYL 480
XX
XX 421 ASVQSKRRKRLRYLHCTNTDNRKYEGDLTYAIVLHNVTKLRLPYPSNKNQVVKYL 480
XX 421 ASVQSKRRKRLRYLHCTNTDNRKYEGDLTYAIVLHNVTKLRLPYPSNKNQVVKYL 480
XX
XX 481 RPLGPHGLSKSVQNLGLTKMVDQTLPLMEKPLRPGSSGLPAFSSFFYIRAKYA 540
XX 481 RPLGPHGLSKSVQNLGLTKMVDQTLPLMEKPLRPGSSGLPAFSSFFYIRAKYA 540
XX
XX 481 RPLGPHGLSKSVQNLGLTKMVDQTLPLMEKPLRPGSSGLPAFSSFFYIRAKYA 540
XX 481 RPLGPHGLSKSVQNLGLTKMVDQTLPLMEKPLRPGSSGLPAFSSFFYIRAKYA 540
XX
XX 541 ACI 543
XX |||

```

```

DB 541 ACI 543
XX
XX RESULT 8
XX ADL16379
XX ID ADL16379 standard; protein; 543 AA.
XX
XX AC ADL16379;
XX
XX DT 06-MAY-2004 (first entry)
XX
XX DE Human heparanase partial protein.
XX
XX KW Human; heparanase; heparanase-dependent cancer; cancer;
XX autoimmune reaction; inflammation; chromosome 4; enzyme.
XX
XX OS Homo sapiens.
XX
XX PN US2003236215-A1.
XX
XX PD 25-DEC-2003.
XX
XX PF 09-JUN-2003; 2003US-00456573.
XX
XX PR 31-AUG-1998; 98WO-US017954.
XX PR 01-MAR-1999; 99US-00258892.
XX PR 08-NOV-1999; 99US-00435739.
XX
XX (INSI-) INSIGHT STRATEGY & MARKETING LTD.
XX (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
XX
XX Pecker I, Vladavsky I, Feinstein E;
XX
XX MPI: 2004-070610/07.
XX
XX PT New antisense oligonucleotide hybridizable with a polynucleotide encoding
XX a polypeptide with heparanase activity, useful for treating diseases such
XX as cancer and autoimmune disorders.
XX
XX PS Claim 3: SEQ ID NO 10; 108bp; English.
XX
XX The invention relates to an antisense oligonucleotide (ASO) comprising a
XX polynucleotide or a polynucleotide analogue of at least 10 bases being
XX hybridizable in vivo, under physiological conditions, with a portion of
XX a polynucleotide strand encoding a polypeptide having heparanase
XX catalytic activity. Also included are a method of in vivo downregulating
XX heparanase activity (comprising administering the ASO in vivo), a method
XX of treating a subject suffering from a pathological condition
XX (characterised by heparanase activity, comprising administering ASO to
XX the subject), a pharmaceutical composition comprising the ASO and a
XX carrier, an antisense nucleic acid construct (comprising a promoter
XX sequence and a polynucleotide sequence directing the synthesis of an
XX antisense RNA sequence of at least 10 bases being hybridizable in vivo,
XX under physiological conditions, with a polynucleotide strand encoding a
XX polypeptide having heparanase catalytic activity), a method of in vivo
XX downregulating heparanase activity (comprising administering in vivo the
XX antisense nucleic acid construct), a pharmaceutical composition
XX comprising the antisense nucleic acid construct and a carrier, and an
XX antisense oligonucleotide comprising a polynucleotide or a polynucleotide
XX analogue of at least 10 bases being hybridizable in vivo, under
XX physiological conditions, with a portion of a polynucleotide strand being
XX characterised by forming at least a portion of an untranslated region
XX (UTR) for a polynucleotide strand encoding a polypeptide having
XX heparanase catalytic activity. The methods and compositions of the
XX present invention are useful for the prevention and/or treatment of
XX diseases or conditions associated with aberrant heparanase activity, such
XX as heparanase-dependent cancer, cancer, autoimmune reaction and
XX inflammation. The gene for human heparanase is located on chromosome 4.
XX
XX The present sequence is a human heparanase protein.
XX
XX SQ Sequence 543 AA;
XX
XX Query Match 100.0%; Score 2842; DB 8; Length 543;

```

Db 241 QLGEDYIQLHKLRLKSTFKNATLYGPDVGQPRRTAKMLKSFLLKAGGEVIDSVTHHHYLL 300
 QY 301 NGRTATREDPLNPDVLDIFISSVQKVPQVVESTRGKKVWLGETSSAYGGAPLLSDTFA 360
 Db 301 NGRTATREDPLNPDVLDIFISSVQKVPQVVESTRGKKVWLGETSSAYGGAPLLSDTFA 360
 QY 361 AGFWMLDKLGLSARMGIEVVMRQVFPGAGNYHLVDENPDLDPDWLSLLFKKLVTGTVLM 420
 Db 361 AGFWMLDKLGLSARMGIEVVMRQVFPGAGNYHLVDENPDLDPDWLSLLFKKLVTGTVLM 420
 QY 421 ASVOGSKRRKRLRVYLHCTNTDNPXYKEGDLTYAINTLHNTYKYLRLPYPSNKQVDKYL 480
 Db 421 ASVOGSKRRKRLRVYLHCTNTDNPXYKEGDLTYAINTLHNTYKYLRLPYPSNKQVDKYL 480
 QY 481 RPLGPHGLSKSVQNLGLTLKXVDDQTLPLMEKPLRPSSSLGLPAFSYSPFVIRAKYA 540
 Db 481 RPLGPHGLSKSVQNLGLTLKXVDDQTLPLMEKPLRPSSSLGLPAFSYSPFVIRAKYA 540
 QY 541 ACI 543
 Db 541 ACI 543

RESULT 6
 ABB07813
 ID ABB07813 standard; protein; 543 AA.

AC ABB07813;
 XX
 DT 03-JUL-2002 (first entry)
 XX
 DE Human heparanase sequence.
 XX
 DE Heparanase; catalytic; cytosolic; antiviral; antibacterial; enzyme;
 KW anti-protocan; neuroprotective; heparin; human.
 XX
 OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT Peptide 1..35
 FT /note= "signal peptide"
 FT Protein 36..543
 FT /note= "mature protein"

XX US2002034810-A1.
 XX PD 21-MAR-2002.
 XX PF 16-AUG-2001; 2001US-00930218.
 XX PR 20-SEP-2000; 2000US-00666390.

XX (INST-) INSIGHT STRATEGY & MARKETING LTD.
 PA Goldsmith O, Pecker I, Vlodavsky I, Michal I, Zcharia E;

XX WPI; 2002-338926/37.
 DR
 XX
 PT

PT Nucleic acid encoding avian and reptile heparanase polypeptide is useful
 PT to treat various heparin-related disorders and the signal peptide is
 PT useful in production of membrane-targeted or secreted recombinant
 PT proteins.
 XX
 PS
 XX

XX Disclosure; Fig 1a; 39pp; English.

CC The invention relates to an isolated avian and reptile nucleic acid,
 CC encoding a polypeptide with heparanase catalytic activity. The signal
 CC peptide of the nucleic acid can be used to express membrane-associated or
 CC secreted proteins in heterologous expression systems. The encoded
 CC polypeptides can be used to prevent tumour angiogenesis, metastasis and
 CC invasion, and to intervene with pathologies associated with impaired
 CC heparin-binding growth factors, cellular responses to heparin-binding
 CC growth factors and cytokines, cell interaction with plasma lipoproteins,

CC cellular susceptibility to viral, protozoa and bacterial infections or
 CC disintegration of neurodegenerative plaques. The present sequence
 CC represents a human heparanase protein sequence used in similarity studies
 XX
 SQ Sequence 543 AA;

Query Match 100.0%; Score 2842; DB 5; Length 543;
 Best Local Similarity 100.0%; Pred. No. 2.56-273;
 Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLRSKPLPPLMLLLGLGRLSPALPRPAQADVDDLPFQEPHLHVSPLSTYT 60
 Db 1 MLRSKPLPPLMLLLGLGRLSPALPRPAQADVDDLPFQEPHLHVSPLSTYT 60
 QY 61 IDANLATPRFLILGSPRLTARGLSPAYLRFQGTDTPLIPDKKSTEEERSYQS 120
 Db 61 IDANLATPRFLILGSPRLTARGLSPAYLRFQGTDTPLIPDKKSTEEERSYQS 120
 QY 121 QVNODICRYGSLPPDVEEKLRLWEPYQOQLLREHYQKKFKNSTYSRSSVDVLYTFPANC 180
 Db 121 QVNODICRYGSLPPDVEEKLRLWEPYQOQLLREHYQKKFKNSTYSRSSVDVLYTFPANC 180
 QY 181 GDLIFGLNALRLRTADLQWNSNAQLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
 Db 181 GDLIFGLNALRLRTADLQWNSNAQLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
 QY 241 QLGEDYIQLHKLRLKSTFKNATLYGPDVGQPRRTAKMLKSFLLKAGGEVIDSVTHHHYLL 300
 Db 241 QLGEDYIQLHKLRLKSTFKNATLYGPDVGQPRRTAKMLKSFLLKAGGEVIDSVTHHHYLL 300
 QY 301 NGRTATREDPLNPDVLDIFISSVQKVPQVVESTRGKKVWLGETSSAYGGAPLLSDTFA 360
 Db 301 NGRTATREDPLNPDVLDIFISSVQKVPQVVESTRGKKVWLGETSSAYGGAPLLSDTFA 360
 QY 361 AGFWMLDKLGLSARMGIEVVMRQVFPGAGNYHLVDENPDLDPDWLSLLFKKLVTGTVLM 420
 Db 361 AGFWMLDKLGLSARMGIEVVMRQVFPGAGNYHLVDENPDLDPDWLSLLFKKLVTGTVLM 420
 QY 421 ASVOGSKRRKRLRVYLHCTNTDNPXYKEGDLTYAINTLHNTYKYLRLPYPSNKQVDKYL 480
 Db 421 ASVOGSKRRKRLRVYLHCTNTDNPXYKEGDLTYAINTLHNTYKYLRLPYPSNKQVDKYL 480
 QY 481 RPLGPHGLSKSVQNLGLTLKXVDDQTLPLMEKPLRPSSSLGLPAFSYSPFVIRAKYA 540
 Db 481 RPLGPHGLSKSVQNLGLTLKXVDDQTLPLMEKPLRPSSSLGLPAFSYSPFVIRAKYA 540
 QY 541 ACI 543
 Db 541 ACI 543

RESULT 7
 ADG88800
 ID ADG88800 standard; protein; 543 AA.

XX ADG88800;
 AC
 XX
 DT 11-MAR-2004 (first entry)
 XX
 DE Human hpa protein.

KW Wound healing; heparanase; ulcer; burn; laceration; surgical incision;
 KW necrosis; pressure wound; diabetic ulcer; angiogenesis; human; therapy.
 XX
 OS Homo sapiens.

XX US2003161823-A1.
 XX 28-AUG-2003.
 XX 14-JAN-2003; 2003US-00341582.
 XX 31-AUG-1998; 98WO-US017954.

PR 01-MAY-1998; 98US-00071739.
 XX (INSI-) INSIGHT STRATEGY & MARKETING LTD.
 PA (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
 PA (FRIE/) FRIEDMAN M M.
 XX
 PI Pecker I, Vlodavsky I, Friedman Y, Perets T;
 XX WPI; 2000-052944/04.
 DR N-PSDB; AA233290.
 XX
 PT Heparanase-specific molecular probes useful for diagnosis and treatment,
 PT e.g. of tumors, and for targeted drug delivery.
 XX
 PS Example; Page 81-82; 90pp; English.
 XX
 CC The present invention describes heparanase-specific molecular probes,
 CC useful for methods of detecting heparanase in situ. The probes and anti-
 CC heparanase antibodies are used to detect or quantify the expression of
 CC heparanase, for diagnosis and monitoring of diseases (especially
 CC metastasis), for treatment of heparanase-associated diseases (e.g.
 CC tumors, (adeno)carcinoma, squamous cell carcinoma, teratocarcinoma,
 CC mesothelioma, melanoma, lymphoma or leukemia, a solid cancer (or its
 CC metastases) derived from liver, prostate, bladder, breast, ovary, cervix,
 CC colon, skin, intestine, stomach, uterus and pancreas, kidney disease,
 CC diabetes and inflammation, haemorrhagic nephritis, nephrotic syndrome,
 CC sepsis and inflammatory or autoimmune disease), for targeted drug
 CC delivery (e.g. of anticancer agents) and as research reagents. The
 CC present sequence represents human heparanase, which is used in the
 CC exemplification of the present invention
 XX
 SQ Sequence 543 AA;
 Query Match 100.0%; Score 2842; DB 3; Length 543;
 Best Local Similarity 100.0%; Pred. No. 2.5e-273;
 Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 541 ACI 543
 DB 541 ACI 543
 RESULT 5
 ID AAY97635
 AA97635 standard; protein; 543 AA.
 XX
 AC AAY97635;
 XX
 DT 20-APR-2001 (first entry)
 XX
 DE Human heparanase protein sequence.
 XX
 KW Heparanase; hnhp1; wound healing; angiogenesis; restenosis; scrape;
 KW atherosclerosis; inflammation; pulmonary disease; Alzheimer's disease;
 KW neurodegenerative disease; Creutzfeldt-Jakob disease; viral infection;
 KW gene therapy; human.
 XX
 OS Homo sapiens.
 XX
 PN MO200100643-AA.
 XX
 PD 04-JAN-2001.
 XX
 PF 19-JUN-2000; 2000WO-IL000358.
 XX
 PR 25-JUN-1999; 99US-0140801P.
 XX
 PA (INSI-) INSIGHT STRATEGY & MARKETING LTD.
 XX
 PI Pecker I, Michael I, Itzhaki H;
 XX WPI; 2001-137930/14.
 DR
 XX
 PT New polynucleotides and polypeptides that are distantly homologous to
 PT heparanase, useful in wound healing, as well as in gene therapy protocols
 PT for angiogenesis, restenosis, atherosclerosis, or inflammation.
 XX
 PS Disclosure; Page 64-65; 67pp; English.
 XX
 CC This sequence represents a heparanase of the invention. The heparanase
 CC DNA and protein sequences are useful in wound healing, angiogenesis,
 CC restenosis, atherosclerosis, inflammation, pulmonary diseases,
 CC neurodegenerative diseases (such as scrape, Alzheimer's disease, and
 CC Creutzfeldt-Jakob disease) or viral infections. The heparanase coding
 CC sequence is particularly useful in gene therapy
 XX
 SQ Sequence 543 AA;
 Query Match 100.0%; Score 2842; DB 4; Length 543;
 Best Local Similarity 100.0%; Pred. No. 2.5e-273;
 Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 361 AGFMWLDKLGTSARMGIEVVMRQVFFGAGNYHLDVENPDLDPYMLSLFKLVGTXYLM 420
 DB 361 AGFMWLDKLGTSARMGIEVVMRQVFFGAGNYHLDVENPDLDPYMLSLFKLVGTXYLM 420
 QY 421 ASVQSKRRKRLVYLLHCTNTDNPVRYEGDLTYAINTLVNTYKRLPYPFPSNKQVDKXYLL 480
 DB 421 ASVQSKRRKRLVYLLHCTNTDNPVRYEGDLTYAINTLVNTYKRLPYPFPSNKQVDKXYLL 480
 QY 481 RPLGPHGLSKSVQNLGLTKWVDQDTLPPLMEKPLRPSSISGLPFAFSYFVIRNAKYA 540
 DB 481 RPLGPHGLSKSVQNLGLTKWVDQDTLPPLMEKPLRPSSISGLPFAFSYFVIRNAKYA 540
 QY 541 ACT 543
 DB 541 ACT 543
 RESULT 3
 AAB08849
 ID AAB08849 standard; protein; 543 AA.
 AC AAB08849;
 DT 15-JAN-2001 (first entry)
 DE Amino acid sequence of a human heparanase polypeptide.
 XX Human; heparanase; gene therapy; tumour; inflammation; autoimmunity;
 KM heparin-binding growth factor; cytokine; neurodegenerative plaque;
 KM wound healing; infection; burn; angiogenesis; restenosis;
 KM atherosclerosis; inflammation; neurodegenerative disease;
 KM Gerstmann-Straussler Syndrome; Creutzfeldt-Jakob disease.
 OS Homo sapiens.
 PN WO200052178-A1.
 XX 08-SEP-2000.
 PD 14-FEB-2000; 2000WO-US003542.
 PF 01-MAR-1999; 99US-00258892.
 PR (INSI-) INSIGHT STRATEGY & MARKETING LTD.
 PA (HADA-) HADASTI MEDICAL RES SERVICES & DEV.
 PA (FRIE/) FRIEDMAN M W.
 PI Pecker I, Vlodaevsky I, Feinstein E;
 XX MPI: 2000-579289/54.
 DR N-PSDB; AAA75051.
 XX New polynucleotides encoding a polypeptide having heparanase activity,
 PT useful in wound healing and in gene therapy, particularly in treating
 PT tumor, inflammation, autoimmunity, neurodegenerative diseases.
 XX Claim 22; Fig 1; 152bp; English.
 XX The present sequence represents a human protein with heparanase catalytic
 CC activity. The heparanase (hpa) polynucleotide is useful in gene therapy,
 CC particularly in treating tumour, inflammation or autoimmunity.
 CC Particularly, the polynucleotide is useful in modulating the
 CC bioavailability of heparin-binding growth factors, cellular responses to
 CC heparin-binding growth factors (e.g. bFGF) and cytokines (e.g.
 CC interleukin (IL)-8), cell interaction with plasma lipoproteins, cellular
 CC susceptibility to certain viral and some bacterial and protozoa
 CC infections, or disintegration of neurodegenerative plaques. The
 CC polynucleotide is also useful in wound healing (e.g. thermal, chemical or
 CC radiation burns), and in the treatment of angiogenesis, restenosis,
 CC atherosclerosis, inflammation, neurodegenerative diseases (Gerstmann-
 CC Straussler Syndrome or Creutzfeldt-Jakob disease), and some viral,
 CC bacterial or protozoa infections

XX SQ Sequence 543 AA;
 Query Match 100.0%; Score 2842; DB 3; Length 543;
 Best Local Similarity 100.0%; Pred. No. 2,56-273;
 Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLRSKPLPPLMLLLGLPLSPGALPRPAQADVDVLDFTQEPHLHVSPLSVT 60
 DB 1 MLRSKPLPPLMLLLGLPLSPGALPRPAQADVDVLDFTQEPHLHVSPLSVT 60
 QY 61 IDANLATPRLILLGSPKRLTARGLSPAYIRFGGTDFLIPPKKSTEEERWYMS 120
 DB 61 IDANLATPRLILLGSPKRLTARGLSPAYIRFGGTDFLIPPKKSTEEERWYMS 120
 QY 121 QVNODICKYGSIPDVEEKLRLFWPQEOQLLREHYOKKFNSTYSSVVDLYTFANCS 180
 DB 121 QVNODICKYGSIPDVEEKLRLFWPQEOQLLREHYOKKFNSTYSSVVDLYTFANCS 180
 QY 181 GDLIFGLNALRTADLQWSSNAQLLDYCSSKGYNISWELGNEBNSFLKKADIFINGS 240
 DB 181 GDLIFGLNALRTADLQWSSNAQLLDYCSSKGYNISWELGNEBNSFLKKADIFINGS 240
 QY 241 QLGEDYIQLHKLIRKSTFKNATLYGPDVQOPRRKTAKMLKSTFLKAGGEYIDSVTHHHYLL 300
 DB 241 QLGEDYIQLHKLIRKSTFKNATLYGPDVQOPRRKTAKMLKSTFLKAGGEYIDSVTHHHYLL 300
 QY 301 NGRTATREDPLNDVLDIFISSVOKVFOVVESTRPCKKVMLGSTSSAYGGAPILSDTPA 360
 DB 301 NGRTATREDPLNDVLDIFISSVOKVFOVVESTRPCKKVMLGSTSSAYGGAPILSDTPA 360
 QY 361 AGFMWLDKLGTSARMGIEVVMRQVFFGAGNYHLDVENPDLDPYMLSLFKLVGTXYLM 420
 DB 361 AGFMWLDKLGTSARMGIEVVMRQVFFGAGNYHLDVENPDLDPYMLSLFKLVGTXYLM 420
 QY 421 ASVQSKRRKRLVYLLHCTNTDNPVRYEGDLTYAINTLVNTYKRLPYPFPSNKQVDKXYLL 480
 DB 421 ASVQSKRRKRLVYLLHCTNTDNPVRYEGDLTYAINTLVNTYKRLPYPFPSNKQVDKXYLL 480
 QY 481 RPLGPHGLSKSVQNLGLTKWVDQDTLPPLMEKPLRPSSISGLPFAFSYFVIRNAKYA 540
 DB 481 RPLGPHGLSKSVQNLGLTKWVDQDTLPPLMEKPLRPSSISGLPFAFSYFVIRNAKYA 540
 QY 541 ACT 543
 DB 541 ACT 543
 RESULT 4
 AAY52990
 ID AAY52990 standard; protein; 543 AA.
 AC AAY52990;
 DT 21-FEB-2000 (first entry)
 DE Human heparanase protein sequence.
 XX Human; heparanase; hpa; diagnosis; therapy; tumour; cytostatic;
 KM antidiabetic; immunomodulatory; anti-inflammatory; nephrotropic;
 KM metastasis; adenocarcinoma; squamous cell carcinoma; teratocarcinoma;
 KM mesothelioma; melanoma; lymphoma; leukemia; cancer; sepsis; diabetes;
 KM inflammation; haemorrhagic nephritis; nephrotic syndrome;
 KM autoimmune disease; anticancer; kidney disease.
 OS Homo sapiens.
 PN WO9957153-A1.
 PD 11-NOV-1999.
 PF 29-APR-1999; 99WO-US009255.

CC conditions such as wound healing, angiogenesis, restenosis,
 CC atherosclerosis, inflammation, neurodegenerative diseases, and viral
 CC infections. Mammalian heparanase can be used to neutralize plasma
 CC heparin, and anti-heparanase antibodies may be applied for
 CC immunodetection and diagnosis of micrometastases, autoimmune lesions, and
 CC renal failure in biopsy specimens, plasma samples, and body fluids. The
 CC present sequence represents human heparanase

XX Sequence 543 AA;

Query Match 100.0%; Score 2842; DB 2; Length 543;

Best Local Similarity 100.0%; Pred. No. 2.5e-273; Indels 0; Gaps 0;

Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLRSKPALPPMLLLGLPLSPGALPRPAQADVDLDFFTQEPHLVSPSLSYT 60
 DB 1 MLRSKPALPPMLLLGLPLSPGALPRPAQADVDLDFFTQEPHLVSPSLSYT 60
 QY 61 IDANLATDPRFLILGSPKRLTARGLSPAYLRFPGTKTDFLIPDKKSTFEERSYQ 120
 DB 61 IDANLATDPRFLILGSPKRLTARGLSPAYLRFPGTKTDFLIPDKKSTFEERSYQ 120
 QY 121 QVNODICKYGSIPDVVEKRLLEWYQEOQLLREHYOKKFKNSTYRSRVDLVYTFANCS 180
 DB 121 QVNODICKYGSIPDVVEKRLLEWYQEOQLLREHYOKKFKNSTYRSRVDLVYTFANCS 180
 QY 181 GDLIFGLNALIRTDLQWNSNAQLLDYCSKGYNISWELGNENPSFLKADIFINGS 240
 DB 181 GDLIFGLNALIRTDLQWNSNAQLLDYCSKGYNISWELGNENPSFLKADIFINGS 240
 QY 241 QUGEDYIQLHKLIRKSTFKNAKLYGPDVGQPRRTAKMLKSLKAGGEVIDSVTHHHYLL 300
 DB 241 QUGEDYIQLHKLIRKSTFKNAKLYGPDVGQPRRTAKMLKSLKAGGEVIDSVTHHHYLL 300
 QY 301 NGRATREDFLNPVDLFISSVQKVFQVVESTRPGKKVWLGETSSAYGGAPLLSDTFA 360
 DB 301 NGRATREDFLNPVDLFISSVQKVFQVVESTRPGKKVWLGETSSAYGGAPLLSDTFA 360
 QY 361 AGFMWLDKLGLSARMGIEVVMQVFFGAGNYHLVDENPDPLPDYMLSLFKLVGKVLAM 420
 DB 361 AGFMWLDKLGLSARMGIEVVMQVFFGAGNYHLVDENPDPLPDYMLSLFKLVGKVLAM 420
 QY 421 ASVQSSKRRKRLRYVHCTNTDNPRIYEGDLTYAINLNAVTKYLRLLPYPSNKQVDKYL 480
 DB 421 ASVQSSKRRKRLRYVHCTNTDNPRIYEGDLTYAINLNAVTKYLRLLPYPSNKQVDKYL 480
 QY 481 RPLGPHGLSKSVQNLGLTLKAVDDOTLPRIMEKPLRPSSSLGLPAFSSFPVIRAKYA 540
 DB 481 RPLGPHGLSKSVQNLGLTLKAVDDOTLPRIMEKPLRPSSSLGLPAFSSFPVIRAKYA 540
 QY 541 ACI 543
 DB 541 ACI 543

RESULT 2
 ID AAY57590
 ID AAY57590 standard; protein; 543 AA.

XX AAY57590;

XX 02-MAR-2000 (first entry)

XX Human heparanase.

XX Human, heparanase; hpa; genetic modification; expression; anticancer;
 XX angiogenesis; anti-angiogenic; antiproliferative; antiviral; antitumour;
 XX anti-atherosclerotic; anti-inflammatory; antineurodegeneration;
 XX heparan sulphate; heparin-binding growth factor; tumour angiogenesis;
 XX metastasis; wound healing; restenosis; atherosclerosis; inflammation;
 XX neurodegeneration; viral infection; cystic fibrosis; cancer; diagnosis;
 XX micrometastasis; autoimmune lesion; kidney failure.

OS Homo sapiens.

XX WO9557244-A1.

XX 11-NOV-1999.

XX 29-APR-1999; 99WO-US009256.

XX 01-MAY-1998; 98US-00071618.

XX 02-MAR-1999; 99US-00260038.

XX (INSI-) INSIGHT STRATEGY & MARKETING LTD.

XX (FRIE/) FRIEDMAN M.W.

XX Ben-Arzi H, Ayal-HersHKovitz M, Yacoby-Zeevi O, Pecker I;

XX Pelag Y, Shlomi Y;

XX WPI; 2000-062144/05.

XX N-PSDB; AA239195.

XX Engineered cells that express recombinant heparanase, useful

XX therapeutically, e.g. for treating angiogenesis and to screen for

XX specific inhibitors, potential anticancer agents.

XX Claim 3; Page 107-109; 118pp; English.

CC The present invention describes genetically modified cells (A) containing
 CC a polynucleotide (I) that encodes a polypeptide with heparanase activity,
 CC and express recombinant heparanase (II). Heparanase cleaves heparan
 CC sulphate (HS) at specific intrachain sites, resulting in release of
 CC heparin-binding growth factors, enzymes and proteins that are sequestered
 CC by HS in basement membranes, extracellular matrix or cell surfaces. It
 CC may also be implicated in tumour angiogenesis and metastases. (II) is
 CC potentially useful in wound healing and for treating angiogenesis,
 CC restenosis, atherosclerosis, inflammation, neurodegeneration, viral
 CC infection and cystic fibrosis. It can also be used to neutralise heparin
 CC (an alternative to protamine) and to screen for specific inhibitors
 CC (potentially useful for treating cancer and metastases). Antibodies
 CC raised against (II) are used for immunodetection and diagnosis of
 CC micrometastases, autoimmune lesions and kidney failure. (A) provide (II)
 CC in large quantities, in a form that is homogeneously processed and
 CC activated/neutralised by a dedicated protease. The present sequence
 CC represents human heparanase

XX Sequence 543 AA;

Query Match 100.0%; Score 2842; DB 3; Length 543;

Best Local Similarity 100.0%; Pred. No. 2.5e-273; Indels 0; Gaps 0;

Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLRSKPALPPMLLLGLPLSPGALPRPAQADVDLDFFTQEPHLVSPSLSYT 60
 DB 1 MLRSKPALPPMLLLGLPLSPGALPRPAQADVDLDFFTQEPHLVSPSLSYT 60
 QY 61 IDANLATDPRFLILGSPKRLTARGLSPAYLRFPGTKTDFLIPDKKSTFEERSYQ 120
 DB 61 IDANLATDPRFLILGSPKRLTARGLSPAYLRFPGTKTDFLIPDKKSTFEERSYQ 120
 QY 121 QVNODICKYGSIPDVVEKRLLEWYQEOQLLREHYOKKFKNSTYRSRVDLVYTFANCS 180
 DB 121 QVNODICKYGSIPDVVEKRLLEWYQEOQLLREHYOKKFKNSTYRSRVDLVYTFANCS 180
 QY 181 GDLIFGLNALIRTDLQWNSNAQLLDYCSKGYNISWELGNENPSFLKADIFINGS 240
 DB 181 GDLIFGLNALIRTDLQWNSNAQLLDYCSKGYNISWELGNENPSFLKADIFINGS 240
 QY 241 QUGEDYIQLHKLIRKSTFKNAKLYGPDVGQPRRTAKMLKSLKAGGEVIDSVTHHHYLL 300
 DB 241 QUGEDYIQLHKLIRKSTFKNAKLYGPDVGQPRRTAKMLKSLKAGGEVIDSVTHHHYLL 300
 QY 301 NGRATREDFLNPVDLFISSVQKVFQVVESTRPGKKVWLGETSSAYGGAPLLSDTFA 360
 DB 301 NGRATREDFLNPVDLFISSVQKVFQVVESTRPGKKVWLGETSSAYGGAPLLSDTFA 360

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2005, 11:09:57 ; Search time 113.5 Seconds
(without alignments)
1850.316 Million cell updates/sec

Title: SEQ2A
Perfect score: 2842
Sequence: 1 MLRSKPALPPIMLLGP.....LPFSYSPVIRNAKVAACI 543

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues
Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A_Geneseq_16Dec04:*
2: geneseqp19808:*
3: geneseqp19908:*
4: geneseqp20008:*
5: geneseqp20018:*
6: geneseqp20028:*
7: geneseqp20038:*
8: geneseqp20048:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2842	100.0	543	2 AAY02345	Aay02345 A human h
2	2842	100.0	543	3 AAY57590	Aay57590 Human hep
3	2842	100.0	543	3 AAB08849	Aab08849 Amino aci
4	2842	100.0	543	3 AAY52990	Aay52990 Human hep
5	2842	100.0	543	4 AAY7635	Aay7635 Human hep
6	2842	100.0	543	5 ABB07813	Abb07813 Human hep
7	2842	100.0	543	7 ADG88800	Adg88800 Human hep
8	2842	100.0	543	7 ADL16379	Adl16379 Human hep
9	2842	100.0	543	8 ADMA8716	Adma8716 Human hpa
10	2842	100.0	543	8 AAY02346	Aay02346 A human h
11	2842	100.0	543	3 AAB08850	Aab08850 Amino aci
12	2842	100.0	543	7 ADG88804	Adg88804 Human hep
13	2842	100.0	543	7 ADL16383	Adl16383 Human hep
14	2842	100.0	543	8 ADMA8720	Adma8720 Human SK-
15	2842	100.0	543	2 AAY17082	Aay17082 Human hep
16	2838	99.9	543	4 AAB86206	Aab86206 Human hep
17	2838	99.9	543	4 ADD18950	Add18950 Human dis
18	2838	99.9	543	8 ADK52086	Adk52086 Human ato
19	2838	99.9	543	8 ADMA8759	Adma8759 Human hpa
20	2838	99.9	543	8 ADNO5074	Adno5074 Antipsoi
21	2838	99.9	543	8 ADNO4902	Adno4902 Antipsoi
22	2838	99.9	543	8 ADQ80372	Adq80372 Hepatana
23	2838	99.9	543	8 ADR88210	Adr88210 Human pre
24	2838	99.9	543	8 ADP25079	Adp25079 PRC Polyp
25	2838	99.9	588	2 AAY30124	Aay30124 A human p

26	2832	99.6	543	8 ADO63831	Ado63831 Human hep
27	2832	99.6	543	8 ADO63823	Ado63823 Human hep
28	2832	99.6	543	8 ADO63832	Ado63832 Human hep
29	2832	99.6	543	8 ADO63822	Ado63822 Human hep
30	2826	99.4	543	4 AAB88361	Aab88361 Human mem
31	2826	99.4	543	8 ADO63824	Ado63824 Human hep
32	2817	99.1	545	6 ABP56822	Abp56822 Human hep
33	2817	99.1	545	7 ADE16012	Adel6012 G-coupled
34	2817	99.1	545	8 ADL93951	Adl93951 Human G-c
35	2764	97.3	530	2 AAY34173	Aay34173 Human pre
36	2737	96.3	532	2 AAY17083	Aay17083 Seq ID No
37	2673.5	94.1	527	5 AAB07815	Abb07815 Chicken s
38	2673.5	94.1	527	7 ABR02018	Abro2018 Chimeric
39	2667.5	93.9	527	8 ADO63825	Ado63825 Chimeric
40	2667.5	93.9	527	8 ADO63826	Ado63826 Chimeric
41	2661.5	93.6	527	8 ADO63827	Ado63827 Chimeric
42	2146	75.5	535	3 AAB08851	Aab08851 A murine
43	2146	75.5	535	5 ABB07811	Abb07811 Mouse hep
44	2146	75.5	535	7 ADG88834	Adg88834 Mouse hpa
45	2146	75.5	535	8 ADL16413	Adl16413 Mouse hep

ALIGNMENTS

RESULT 1	
AAY02345	
ID AAY02345 standard; protein; 543 AA.	
AC AAY02345;	
DT 09-JUL-1999 (first entry)	
DE A human heparanase protein.	
HE	
HE	
KW Heparanase; hp; modulator; heparin-binding growth factor;	
KW cellular response; cytokine; cell interaction; plasma lipoprotein;	
KW cellular susceptibility; infection; disintegration;	
KW neurodegenerative plaque; wound healing; angiogenesis; restenosis;	
KW atherosclerosis; inflammation; neurodegenerative disease; neutralise;	
KW plasma heparin; micrometastasis; autoimmune lesion; renal failure.	
XX	
XX Homo sapiens.	
OS	
XX	
XX WO9911798-A1.	
PN 11-MAR-1999.	
XX	
XX 31-AUG-1998; 98WO-US017954.	
PF	
XX 02-SEP-1997; 97US-00922170.	
PR	
XX 02-JUL-1998; 98US-00109386.	
PA	
PA (INSI-) INSIGHT STRATEGY & MARKETING LTD.	
PA (HADA-) HADASIT MEDICAL RES SERVICES & DEV.	
PA (FRIE/) FRIEDMAN M M.	
XX	
XX Pecker I, Vlodavsky I, Feinstein E;	
PI	
XX	
XX WPI, 1999-302255/25.	
DR	
XX N-PSDB; AAX35648.	
PT	
XX New human polynucleotide useful for treating angiogenesis, restenosis,	
PT and inflammation.	
XX	
XX Claim 6, Fig 1; 63pp; English.	
PS	
XX The specification describes a polypeptide having heparanase (hp)	
XX activity. The recombinant protein is used as a modulator of heparin-	
CC binding growth factors, cellular responses to heparin-binding growth	
CC factors and cytokines, cell interaction with plasma lipoproteins,	
CC cellular susceptibility to viral, protozoal and bacterial infections or	
CC disintegration of neurodegenerative plaques. Heparanase may be useful for	

This Page Blank (uspto)

```

QY 106 PKKESTFEERSYQWQVNOIDICKYGSIPDVEBKLRLMPYQEQOLLREHYQOKKFNSTY 165
DB 100 -SKENI-----PISCHNCSYKSPKSLQ--LIEKPC-----KHKKFLPFLM 140
QY 166 SSSVDVLYTPANCGLDILIFGLNALLRTRADLOWNSNAQOLLIDYSSSGYNI SWELGNE 225
DB 141 TGNENWQIJDPCRKTMLKLLFSLNAMLRD--NKGWNEKNARELIEFSEKHQVYIDWQLGNE 199
QY 226 PMSFLKKADIFINGSLGEDIQLHKLRLKSTPFNAKLGYPDVGOP---RRKTAKMLKSF 282
DB 200 PMSFGHVESYTPQILADQFEKRLKLNHNGYRSHLIVGPDTRTPQPRPELKTMIFF 259
QY 283 LKXAGEVIDSVTWHYYLNGRTATREDPLNPVDLIDFISVQKPFQVESTPCKKV--W 340
DB 260 LQNGSHYIVNRSHQYLLNSKTAKLEDFNWPEFFDL---RQILEMQQTKYXNI PMW 316
QY 341 LGETSAYGGGAPLSDTPAAGFMWLDKLGSLARMGIEVVMROVFFGAGNYHLVDENPDP 400
DB 317 LSETSSYGGGAPGLSNTYAGSPLMIDKLGSLAKYNI STVIRQSFYG--GYSLVDENLKP 375
QY 401 LFDYWLSTLFFKLVGKVLMAVQGSRRKRLRYLHCTNTDNPRYKE--GDLLTYAIN-- 456
DB 376 LFDWMISVLKYKLVGNKVL--QVQNCGRFQRYLHCTNR--KYNDTSAYTLVYGVNLE 430
QY 457 -----LHN-----VTXYLRLLPYPSNKOVDKYLRLPLGPHGLSKSVQVNLGLTL 500
DB 431 MAKARFPLNGTALHGDDLIHEYL--ISAPSNRK-----SKTLLNGWPL 474
QY 501 KXAVDQTLPLMEKPLRPGSSLGLPAFYSFFPIYIRAKYAAAC 542
DB 475 YV--ESNLHNLRPNIHRYGRVYSLPPYSIGFWIKTKTSITVC 514

RESULT 15
Q9SDAI PRELIMINARY; PRT; 521 AA.
ID Q9SDAI;
AC Q9SDAI;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein F13G24.30.
GN Name=F13G24.30.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Van Der Schueren J., Chuang Y.J., Voet M., Robben J.,
RA Volckaert G., Bancroft I., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL133421; CAB62595.1; -.
DR PIR; T45608; T45608.
DR InterPro; IPR005199; Glyco_hydro_79N.
DR InterPro; IPR001254; Peptidase_S1.
DR Pfam; PF03662; Glyco_hydro_79n; 1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 521 AA; 57831 MW; 07DB664A4B305CC2 CRC64;

```

```

Query Match 14.6%; Score 416; DB 2; Length 521;
Best Local Similarity 29.2%; Pred. No. 1.6e-22;
Matches 154; Conservative 68; Mismatches 184; Indels 122; Gaps 24;

```

```

QY 75 LGSPPKRLTARGSLPAYLRFGCTKIDPLFDPKKESTFEERSYQWQVNOIDICKYGSIPP 134
DB 55 LTRPLLTAKAKPKPIRIGSLQDOVIVYDGNLKT-----PCR----- 94

```

```

QY 135 DVEBKRLRLMPYQEQOLLREHYQOKKFNKS--TYSRSSV-----DLYTFPANCGLDILIF 186
DB 95 -----PQOKM-----NSGLRGFSKGLCHMRKMDLNLFLATGAVYTF 132
QY 187 GLNALIRTRDLO-----WSSNAQOLLIDYSSSGYNI--SWELGNEPMSFLKKADIFIN 238
DB 133 GLNALRGRHKLKGKAGWAMDHINQDFLNTYVSKGYVIDSWFEGNELSG--SGVGSVYS 190
QY 239 GSQLEDYQLHKLRLKSTPFNAKLGYPDVGOP---RRKTAKMLKSFLLKAGEVIDSV 293
DB 191 AELYGKDLIVLKDVIK--VYKNSWLHKPLVAPGGEYEQGWTKLLEI---SGPSVVDV 246
QY 294 TWHYYLNGRT--ATREDFLNPVDLIDFISVQKPF---QVESTPCKKYLWLGSTSA 347
DB 247 THHIVYLGSGNDPALVKKIMDPS---YLSQVSKTFKDVNQIIOEHGPWASPWGSGGA 302
QY 348 YGGGAPLSDTPAAGFMWLDKLGSLARMGIEVVMROVFFGAGNYHLVD--NPDPLDYVL 406
DB 303 YNSGGRHVSDFITDSFWYLDOLGMSARHNTKYCRQTLVG--GFYGLLEKGTVPNPDDYYS 361
QY 407 SLFFKLVGTKVLMAVQGSRRKRLRYLHCTNTDNPRYKEGDLTYAINLHNVTKYL-- 464
DB 362 ALIMHRLMKGVLAVOTDGP--QLRIVYAHCSK-----GRAGVTLLINLSQSDFTYS 413
QY 465 -----RLPYPS---NKQVDKYLRL--LGPHG--LLSKSVQL 495
DB 414 VNSGINVVLNMAESRRKKSLDLTLKRPFSWIGSKASPDGYLNREYHLTPENGVLRSKTMVL 473
QY 496 NGTILTKVDDQTLPLMEKPLRP--GSSLGLPAFYSFFPIYIRAKYAAAC 542
DB 474 NGRSLKPTATGDI PSL--EPVLRSVNSPLNVLPLMSFIVLPNPDASAC 520

```

Search completed: March 23, 2005, 11:29:33
 Job time : 112.5 secs

```

Db 78 FLSTLDPSIINH-CMLDPLSSKRLVTLARGSPFLRFGGKRTFLQONLRNPKSRG 136
Qy 112 FEERSYWGQVNOQDI-----CKYGSIPRVEEKLRLMEPYOQL-LLRHHYK 158
Db 137 GGGPRTYLNKYEDDILRSVVALDKQKCKIAQ-HFVHMLQREKKAQHMLVLEKQF-- 193
Qy 159 KFKNSTYRSRSVAVLYTFPANCGLDLIFGLNALRLTADLQWSSNAQLLDYCSKGYNI 218
Db 194 ---SNTYS-----NIL- 202
Qy 219 SWELGNEPNSFLKADIFINGSQLEDYIQLHKLK-R-STFKNAKLYGPDVQPRKTKAK 277
Db 203 -----TEPNNRYRTMGRAVANGSOLGKDYIQLKSLQPIRIRYGRASLYGNIGRPRKNVIA 257
Qy 278 MLKSPFKAGGEVYIDSVTHHHYLLNGRTATREDFLNPDVLDIFISSVQKRFQVVESTRPK 337
Db 238 LLDGFMKAVASTVDATWQHCHYIDGKVVVWMPFLKTRLDLTDLSQIRKIQKVNTYTPGK 317
Qy 338 KVMLGETSAYGGAFLSDTFAGFMWLDKLSARWGEVWVRQVFFGAGNYHLVDEN 397
Db 318 KIMLBSGVVTSAGTNNISDSYAAGFLWNTLGLMLANGIDVIRHSFFDHGYNHLVDQN 377
Qy 338 FPPLEDYMLSLFKLVGTKVTLMASVQSKRR-----KLRYLHCTNTDNPYKEG 448
Db 378 FNPPLDYMLSLFKLVGTKVTLMASVQSKRR-----KLRYLHCTNTDNPYKEG 437
Qy 449 DLTVAINHANYTKYLRLEFPFNSKQVNDKYLRPLGPHGLSKSVOLNGLTKWDDOTL 508
Db 438 STLTITLNHRSRKKIKLGTLRDLKHVHLLQPGQEGLSKSVQVLANGQPLWVDDGTL 497
Qy 509 PLMEKPLRPGSSGLPAPSVSFVIRNAKVAAC 542
Db 498 PELKRPRLRAGRTLVIPVYTMGFVVKVNNALAC 531

RESULT 13
Q9HB39 PRELIMINARY; PRT; 480 AA.
AC Q9HB39;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Heparanase-like protein HPpaz.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxId=9606;
(1)
RP SEQUENCE FROM N.A.
RX MEDLINE=20483645; PubMed=11027606; DOI=10.1006/dbrc.2000.3586;
RA McKenzie E., Tyson K., Stamps A., Smith P., Turner P., Barry R.,
RA Hirock M., Patel S., Barry E., Stuberfeld C., Terrett J., Page M.;
RT "Cloning and expression profiling of Hpa2, a novel mammalian
RT heparanase family member.";
RL Biochem. Biophys. Res. Commun. 276:1170-1177 (2000).
RN
RN
RN
RP SEQUENCE FROM N.A.
RA McKenzie E.A., Tyson K., Stamps A.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF282885; AAG23421.1; -.
DR PIR; JC7506; JC7506.
DR InterPro; IPR005199; Glyco_hydro_79N.
DR Pfam; PF03662; Glyco_hydro_79n; I.
SQ SEQUENCE 480 AA; 5390 MW; F75F89F67AC1PF83 CRC64;

```

```

Query Match 31.6%; Score 897.5; DB 2; Length 480;
Best Local Similarity 36.0%; Pred. No. 9, 4e-59;
Matches 202; Conservative 74; Mismatches 146; Indels 139; Gaps 9;
20 PLGPIPSGL-----PRPA-----QAQVVDLDFTOPLHVS 55
Db 18 PPACIAPGALVALLLHSLSSQAGDRRLPVDRAAGLKEKTLILDVSTKNPRTV 77

```

```

Qy 56 FLSTIDANLATDPPFLILLGSPKRTTLARGSPAYLRFGGTKDPLIF----DPKEST 111
Db 78 FLSTLDPSIINH-CMLDPLSSKRLVTLARGSPFLRFGGKRTFLQONLRNPKSRG 135
Qy 112 FEERSYWGQVNOQDI-----CKYGSIPRVEEKLRLMEPYOQL-LLRHHYK 171
Db 136 -----GGRGPD-----YLNKYE- 148
Qy 172 VLYTFPANCGLDLIFGLNALRLTADLQWSSNAQLLDYCSKGYNISWELGNEPNSFLK 231
Db 149 -----DEPNNRYRT 156
Qy 232 KADIFINGSQLEDYIQLHKLK-R-STFKNAKLYGPDVQPRKTKAKMLKSPFKAGGEVY 290
Db 157 MHGRAVNSQLOKDYIQLKSLQPIRIRYGRASLYGNIGRPRKNVIALLDGFMKAVASTV 216
Qy 291 DSVTHHHYLLNGRTATREDFLNPDVLDIFISSVQKRFQVVESTRPKKVMLGETSAYG 350
Db 217 DAVTWQHCHYIDGRVVKWMPFLKTRLDLTDLSQIRKIQKVNTYTPGKKIMLEGVTTAG 276
Qy 351 GAPLSDTFPAGFMWLDKLSARWGEVWVRQVFFGAGNYHLVDENPDYMLSLIF 410
Db 277 GTNNISDSYAAGFLWNTLGLMLANGIDVIRHSFFDHGYNHLVDQNFNPDPYMLSLY 336
Qy 411 KLVGTKVTLMASVQSKRR-----KLRYLHCTNTDNPYKEGDLTVAINHANYT 461
Db 337 KRLIPKTLAVHVAAGLQKRRPRGRVIRDKLRIYACHTHHNNHYRGSITLPIULHRSR 396
Qy 462 KYLRLEFPFNSKQVNDKYLRPLGPHGLSKSVOLNGLTKWDDOTLPLMEKPLRPGSS 521
Db 397 KKIKLAGTLRDLKHVHLLQPGQEGLSKSVQVLANGQPLWVDDGTLPELKRPRLRAGRT 456
Qy 522 LGPAPSVSFVIRNAKVAAC 542
Db 457 LVIPVYTMGFVVKVNNALAC 477

RESULT 14
Q8T108 PRELIMINARY; PRT; 515 AA.
AC Q8T108;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Heparanase-like protein.
OC Eukaryota; Metazoa; Chordata; Hexapoda; Insecta; Pterygota;
OC Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC Bombycidae; Bombyx.
NCBI_TaxId=7091;
(1)
RP SEQUENCE FROM N.A.
RC STRAIN=ps0; TISSUE=posterior silkgland;
RA Koike Y., Mita K., Suzuki M.G., Maeda S., Abe H., Osoegawa K.,
RA deJong P.J., Shinada T.;
RT "Genomic sequence of a 320-kb segment of the Z chromosome of Bombyx
RT mori containing a kettin ortholog.";
RL Mol. Genet. Genomics 269:137-149 (2003).
DR EMBL; AB079860; BAB85191.1; -.
DR EMBL; AB090307; BAC10612.1; -.
DR InterPro; IPR005199; Glyco_hydro_79N.
DR Pfam; PF03662; Glyco_hydro_79n; I.
SQ SEQUENCE 515 AA; 59769 MW; FB8100AB6EDDADB CRC64;

```

```

Query Match 24.5%; Score 696; DB 2; Length 515;
Best Local Similarity 35.1%; Pred. No. 1, 4e-43;
Matches 183; Conservative 83; Mismatches 182; Indels 74; Gaps 18;
46 QEPHLVSPSTLSTIDANLATDPPFLILLGSPKRTTLARGSPAYLRFGGTKDPLIF 105
Db 42 QEDIKLISEDFLSGID-TIEIENYNRINYSDTRLRELAALSPARLRLGGTMSERLIF- 99

```

```

QY 20 PLGLPSGAL-----PRPA-----QAQDVVDLDFPTQEPHLVSPS 55
| : : : : :
DB 18 PACACLAGALYALALLHLSSQAGDRRLPVDRAAGLKEKTLILLDVSTKRVRTVNEV 77
| : : : : :
QY 56 FLSTVITDANLADPPRLILGSPKRTTLARGSPAYLRFGGKTDPILF---DPKKEST 111
| : : : : :
DB 78 FLSTLDPSIHD-GWLDPLSSKRLVTLARGSPAYLRFGGKTDPILFQONLRNPAKSRG 136
| : : : : :
QY 112 PEERSYQWQVNDI-----CKYGSIPRVEEKLRLWYEQQL-LIREHYOK 158
| : : : : :
DB 137 GGGPRTYLYKNYEDDVRSDVALDKQCKCIAQ-HPDVMELOREKRAQNHVLLKEQFEN 195
| : : : : :
QY 159 KFKNSTYSSSVYDVLTYTPANCGLDIFGLNALRTADLQWNSNAQLLLDYCSSKXNYI 218
| : : : : :
DB 196 TVSNILITRSLDKLYNSADCSGLHLIFALNLRPNNSMSSSLTKYSASKKXNYI 255
| : : : : :
QY 219 SWELGNEPNSFLKKADIFINGSQLGSDYIQLHKLRLK-STFKNAKLYGSDVQPRKTKX 277
| : : : : :
DB 256 SWELGNEPNRYRTMHGRAVNGSQLGSDYIQLKSLQPIRYSRSLYGNIRPRKNVIA 315
| : : : : :
QY 278 MKSFLKAGGEYIDSTYTHHYLNGRTATREDPLNDVLDIFISSVQKVOYVESTRPGK 337
| : : : : :
DB 316 LLDGFEKAVGASTVDATWQHCYIDGRVVKVMDPLKTRLLDITLSDQIRKIQKVNTYTPGK 375
| : : : : :
QY 338 KYMLGRTSSAYGGAFLSDTPAAGFWMLDKGLSARMGIEVVMRQVFFGAGNYHLVDEN 397
| : : : : :
DB 376 KTWLEGVVTTAGGTNNLSDSYAGFLMNTLGMLANOGIDVYIRHSFDDHGNHLVDQN 435
| : : : : :
QY 398 FPPPLPYWLSLFLKLVGTKVMASVQSKRR-----KLRYVLHCTNDNPKYKKG 448
| : : : : :
DB 436 FNPPLPYWLSLFLKLVGTKVMASVQSKRR-----KLRYVLHCTNDNPKYKKG 495
| : : : : :
QY 449 DTLVYINLHNTKYLRPYPFNSKQVDKYLRLPGLPHGLSKSVOLNGLTKMVDQTL 508
| : : : : :
DB 496 SITLFTIINHRSRKKIKLAGTLRDKLVHGYLLQPYQGEGLKSKSVOLNGLTKMVDQTL 555
| : : : : :
QY 509 PPLMEKPLRPGSGGLPAFSYFFVIRNAKVAAC 542
| : : : : :
DB 556 PELKPRPLRAGRTLVPVPMGFVKVKNVALAC 589
| : : : : :

```

```

RESULT 11
Q9WMQ1 PRELIMINARY; PRT; 548 AA.
AC Q9WMQ1;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hepatanase 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ299720; CAC82492.1; -
DR Pfam; PF03662; Glyco_hydro_79n; 1.
SQ SEQUENCE 548 AA; 61771 MW; B8986303FC73A60A CRC64;

```

```

Query Match 35.7%; Score 1015.5; DB 2; Length 548;
Best Local Similarity 41.8%; Pred. No. 1.6e-67;
Matches 224; Conservative 79; Mismatches 180; Indels 53; Gaps 9;
QY 8 ALPPLMLLGLGLPSGAL-----PRPA-----QAQDVVDLDF 43
| : : : : :

```

```

DB 6 APPEAMLSNSBRPACLAGALYALALLHLSSQAGDRRLPVDRAAGLKEKTLILLDV 65
| : : : : :
QY 44 FTOEPLHVSBSFSLVITDANLADPPRLILGSPKRTTLARGSPAYLRFGGKTDPIL 103
| : : : : :
DB 66 STKNVRVRYNENFSLQDPSIHD-GWLDPLSSKRLVTLARGSPAYLRFGGKTDPILQ 124
| : : : : :
QY 104 F---DPKKESTFEERSYQWQVNDI-----CKYGSIPRVEEKLRLWYEQQL 147
| : : : : :
DB 125 FQNLNPNPAKSRGPGPDYLYKNYEDDVRSDVALDKQCKCIAQ-HPDVMELOREKRAQ 183
| : : : : :
QY 148 EQL-LIREHYOKKFKNSTYSSSVYDVLTYTPANCGLDIFGLNALRTADLQWNSNAQL 206
| : : : : :
DB 194 MEVLVLEKQFNSVYNNLILTRASLDKLYNSADCSGLHLIFALNLRPNNSMSSSLAS 243
| : : : : :
QY 207 LLDYCSSKXNYISWELGNEPNSFLKKADIFINGSQLGSDYIQLHKLRLK-STFKNAKLYG 265
| : : : : :
DB 244 LKYSASAKKXNYISWELGNEPNRYRTMHGRAVNGSQLGSDYIQLKSLQPIRYSRSLYGN 303
| : : : : :
QY 266 PIVGQPRRKTKAMLSFLKAGGEYIDSTYTHHYLNGRTATREDPLNDVLDIFISSVQK 325
| : : : : :
DB 304 PNIGRPRKRVIALLDGFMKVGASTVDATWQHCYIDGRVVKVMDPLKTRLLDITLSDQIRK 363
| : : : : :
QY 326 VQGVNSTRPGKRWMLGRTSSAYGGAFLSDTPAAGFWMLDKGLSARMGIEVVMRQV 385
| : : : : :
DB 364 IOKVNTYTPGKRWMLGRTSSAYGGAFLSDTPAAGFWMLDKGLSARMGIEVVMRQV 423
| : : : : :
QY 386 FGAGNYHLVDENFDPDYWLSLFLKLVGTKVMASVQSKRR-----KLRYVLH 436
| : : : : :
DB 424 PDHGNHLVDQNFNPLPDYWLSLFLKLVGTKVMASVQSKRR-----KLRYVLH 483
| : : : : :
QY 437 CTNTDNPYKESDITLVYINLHNTKYLRPYPFNSKQVDKYLRLPGLPHGLSKSVOLNGL 492
| : : : : :
DB 484 CTNHNNHNVYRGSITLFIINLHRSRKKIKLAGTLRDKLVHGYLLQPYQGEGLKSKSVOL 539
| : : : : :

```

```

RESULT 12
Q9HB38 PRELIMINARY; PRT; 534 AA.
AC Q9HB38;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hepatanase-like protein HP42b.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC MEDLINE=20483645; PubMed=11027606; DOI=10.1006/birc.2000.3586;
RA McKenzie E., Tyson K., Stamps A., Smith P., Turner P., Barry R.,
Hirocock M., Patel S., Barry E., Stubbelfield C., Terrett J., Page M.;
RT "Cloning and expression profiling of Hpa2, a novel mammalian
hepatanase family member";
RL Biochem. Biophys. Res. Commun. 276:1170-1177(2000).
RN (2)
RP SEQUENCE FROM N.A.
RA McKenzie E.A., Tyson K., Stamps A.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF282886; AAC23422.1; -
DR InterPro; IPR005199; Glyco_hydro_79n; 1.
DR Pfam; PF03662; Glyco_hydro_79n; 1.
SQ SEQUENCE 534 AA; 60063 MW; C3DB5E90C8338C4 CRC64;

```

```

Query Match 33.0%; Score 936.5; DB 2; Length 534;
Best Local Similarity 37.8%; Pred. No. 1.3e-61;
Matches 217; Conservative 78; Mismatches 168; Indels 111; Gaps 12;
QY 20 PLGLPSGAL-----PRPA-----QAQDVVDLDFPTQEPHLVSPS 55
| : : : : :
DB 18 PACACLAGALYALALLHLSSQAGDRRLPVDRAAGLKEKTLILLDVSTKRVRTVNEV 77
| : : : : :
QY 56 FLSTVITDANLADPPRLILGSPKRTTLARGSPAYLRFGGKTDPILF---DPKKEST 111
| : : : : :

```

Matches	320: Conservative	87: Mismatches	114: Indels	11: Gaps	3
Qy	13 LMLLLGLPLGSLPGALPRPAQADVDLDEFTQEPHLVSPSPSLVTTIDANLATDPRFL	72			
Db	2 LVLLLLVLLLVLP-----RRTAEIQGLRPIGAVSPAFSLTLDLSIADPFRFV	52			
Qy	73 ILLSPKRLTARAGLSPAYLRFGKTDPLFDPKKESEFEERSRWQGVNDICKYGS	132			
Db	53 ALNRHPKLTHTAASGLSPGLRGGSTDPFLFNPKNKDSWEKVLSEFQA-FDVEAWPMS	111			
Qy	133 PPDEVEKRLRLMPYEOQLRLREHYOKKFKFNSTYSRSSVDVLTFFANCGGLDLIFGLNALL	192			
Db	112 FAVVPKLLLTQWPLEOKKLLLAHSHKSKKHNTTITSTIDILHTFASSSGFRVLPGLNALL	171			
Qy	193 RTADLDONNSVNAQLLDLYCSKSGCNVIMWELGNENPSFLKKADIFINGSQLGEDIYQLHLK	252			
Db	172 RRAGIQMSSNAKQLLGLYCAQGSYNIIMWELGNENPSFRKSGICIDGFOLGDPFHLRQL	231			
Qy	253 L-RKSTFGNAKLYGDPGDPRRKTKAMKSPFKAGAGEVDSVTMHYYLNGRTATREPL	311			
Db	232 LSGHPLYHNAELYGIDVQGPFRKHTOHLRSPFKSGKALIDSTWHYHYNGSARREDPL	291			
Qy	312 NPVDLIDFISSVQKVFQVVESTRPCKYWLGETSSAYGGABLLSDTPAAGFWMLDKGL	371			
Db	292 SPEVLDSFATIAHDVLGIVEATVPCKKTYLGETSGAYGGAPQLSNTYVAGFWMLDKGL	351			
Qy	372 SARMGIEIVMNOVFPGAGNYHVDENPDLPRYWSLLEFKLVGKTYLMAVSQSKRRKL	431			
Db	352 AARRIDIVMNOVSGAGSYHLVDGFKRPLPRYWSLLEKRYLTVGRVYQAQSVQADARRP	411			
Qy	432 RYVLIHCTNDNPRYREGDLYVAIMHWTKYLRYPFSPNKQVQKTYLRLPGPPGLLSK	491			
Db	412 RYVLIHCTNPRHPKYREGDVTLPALNLSNVTQSLQPKOLMSKSVQYLLPHGKDSILSR	471			
Qy	492 SVQLNGLTLLKAVDDQTLPPLEKPLRPGSSGLGPAFYSFFVYIRNAKYAACT	543			
Db	472 EVOJNGRLLOWVDDETLPALHEMALAPGSTGLPAPFSYGFVYIRNAKIACI	523			
RESULT 9					
Q9HB37	PRELIMINARY;	PRT;	592 AA.		
AC	Q9HB37;				
DT	01-MAR-2001 (Tremblrel. 16, Created)				
DT	01-MAR-2001 (Tremblrel. 16, Last sequence update)				
DT	01-JUN-2003 (Tremblrel. 24, Last annotation update)				
DE	Heparanase-like protein Hpa2c.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Euteria; Primates; Carnivora; Homiidae; Homo.				
OX	NCBI_Taxid=9606;				
NP	SEQUENCE FROM N.A.				
RA	MEDLINE=20483445; PubMed=11027606; DOI=10.1006/birc.2000.3586;				
RA	McKenzie E., Tyson K., Stamps A.,				
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AF282887; AAG23423.1; -				
DR	GO; GO:0005622; C:intracellular; TAS.				
DR	GO; GO:0030305; F:heparanase activity; TAS.				
DR	InterPro; IPR005199; Glyco_hydro_79n.				
DR	PIfam; PF03662; Glyco_hydro_79n.1.				
QO	Sequence 592 AA; 66580 MW; 95C384AD9A74258B CRC64;				

```

OY 20 PLGSRBGL-----PRP-----QADVDVDLPEFOELHSPS 55
Db 18 PPAICLPGALVYALLLHLSLSSQADRRPLPVPDRAAGKEKTLILLVSTKNPVRITNEN 77
OY 56 FLVSTIDANLADPREFLILGSPKJLTARGLSPAYLRFQGTCTDPLIF----DPKKEST 111
Db 78 FLBLQDPPSLINH-GWDLFLSSKRLVTLARGLSPALFPGSKRDPFQFQNLNPAASRG 136
OY 112 PEERSYQOVNODI-----CKYGSIPDVEKJLREWPYOBOL-LIREHYOK 156
Db 137 GPGPDYYLKNYEDDIVRSDVALDKQKCKIAQ-HPDVMLLELQREKAQMHLVLLKEQFSN 195
OY 159 KFKNRYTSSSSVDVLYTFANCSCGDLILGIALALTADLCWNSNAOLLDDYCSKRYNI 218
Db 196 TYNLLILTLARSIDKLYNFADSCGHLIFALNALRPNPNNSJALSLSLTSASKRYNI 255
OY 219 SWEIGNEPNSFLKKADIFINGSQLGEDIYOLHKLRLK-STFKNAKLYGPDVQPRRTAK 277
Db 256 SWEIGNEPNNRYRMHGRANVSQLOKDYIOLKSLLOPRIRYSRSLYGPNIGRKRNVIA 315
OY 278 MLKSLFKAAGEVIDSVTWHYYLNGRTATREDFLNPDVLDIFISSVOKVPOVESTREPK 337
Db 316 ILDGFPMKVAGSTVDATVWTOHCYIDRGVVKWDFLKTRLLDITLSDQIRIKIQVWVTYPRGK 375
OY 338 KWLJGESSAYGGGAPLSDTPFAAGFMWLDLGLSARNGIEVWVRQVFPAGNHYLDEN 397
Db 376 KIMLEGVYTSAGCTNNLSDSYAAGFLWMLNTLGMLANQIDVYLRHSFPHDGYNHVLDON 435
OY 398 FDPDLPMVLSLFEKLVGFKVLMASVQSSKKR-----KLAVYLAHCTNVDNPRYKEG 448
Db 436 FNEPLPDWLSLTKRLIGPKVLAVNAGLOKRPGRVIRPKLKIYAHCTNHHNNHYVRG 495
OY 449 DLTLYAINLHNVTKYLRLPYEPSNKQVDYKYLRLPGPHLLSKSVQNLGLTLKXVDDOTL 508
Db 496 SITLFTIINLHRSKKIKLAGLRLDKTLHYQYLLQFYGGLKSKSVQNLQGEVLVWDDGTL 555
OY 509 PPLMEKPLRPGSSLGLPASYSFFYIRAKYAAAC 542
Db 556 PELKPRPLRAGRTLVLPVYTMGFPPVAKVNALAC 589

```

```

RESULT 10
Q8WMQ2                                PRELIMINARY;          PRT;          592 AA.
ID      Q8WMQ2
AC      Q8WMQ2;
DT      01-MAR-2002 (TREMBLrel. 20, Created)
DT      01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT      01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE      Heparataae 2.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Prostate;
RA      Legoux P., Legoux R., O'Brien D., Salome M.;
RL      Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Prostate;
RA      Peseque Safontas B., J.O.P.S.;
RL      Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AJ299719; CAC82491.1; -.
DR      Genew; HGNC:18374; HPS82.
DR      Pfam; PF03662; Glyco_hydro_79n; 1.
SQ      SEQUENCE 592 AA; 66520 MW; 9478841FEACD558B CRC64;

Query Match      40.3%; Score 1146.5; DB 2; Length 592;
Best Local Similarity 43.4%; Pred. NO. 2.5e-77;
Matches 249; Conservative 82; Mismatches 190; Indels 53; Gaps 9;

```

RP SEQUENCE FROM N.A.
RA Huleit M.D., Wang J., Hornby J.R., Freeman C., Pagler E., McHenry J.,
RA Parish C.R.;
RA Submitted (Mar-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF359508; AA015189.1; -
DR InterPro: IPR005199; Glyco_hydro_79N.
DR Pfam: PF03662; Glyco_hydro_79n; 1.
DR SEQUENCE 536 AA; 60479 MW; C434E04CF536E4D CRC64;
Query Match 75.1%; Score 2135; DB 2; Length 536;
Best Local Similarity 76.1%; Pred. No. 1.2e-151;
Matches 407; Conservative 50; Mismatches 78; Indels 0; Gaps 0;
QY 9 LPPMLLLGLPGLSPGALPPPAQADVVLDLFFTOEPHLVSPSLVTIDANLATD 68
DB 2 LRPILLMLMKRLALITGTPAGTAHPKDVVDLEFYTKRLFGQSPSLSTIDASLATD 61
QY 69 PRFLILGSPKRLTLARGSPAYLRFPGTKTDFLIFDPKKESTFEERSYWGQVQNDICK 128
DB 62 PRFLTLGSPRLRALARGSPAYLRFPGTKTDFLIFDPKKEPTSEERSYWGQVQNDICG 121
QY 129 YGSIPTDVEKRLLEMPYQOQLLRHNYQKKFKNSTYSSSDVLYTFANCGDLIFGL 188
DB 122 SERVSADVLRKLOMWPQOELLRLREOYQREPKNSTYSSSDVLYTFANCGDLIFGL 181
QY 189 NALLPTADLQNNSSNAQOLLIDYCSKGYNISWELGNEPNSFLKADIFINGSOUGEYIQ 248
DB 182 NALLPTADLQNNSSNAQOLLIDYCSKGYNISWELGNEPNSFLKADIFINGSOUGEYIQ 241
QY 249 LHKLLRKSTFKNAKLYGPDVQPRRKTAQMLKSLFKAGEVIDSVTWHYYLNGRTATRE 308
DB 242 LHKLLQKSAFQNAKLYGPDIGQPRGKTIVKLRSFLKAGEVIDSLTWHYYLNGRTATRE 301
QY 309 DFLNPDVLDLFISSVQKLYQVYESTRRPKKVMLGFTSSAYGGAPLSDTPAAGFMWLDK 368
DB 302 DFLSSVDVLDLFISSVQKLYQVYESTRRPKKVMLGFTSSAYGGAPLSDTPAAGFMWLDK 361
QY 369 LGLSARMGIEVVMROVFFGAGNYHLVDENFPLPDYWLSTLFFKLVGTVMASVOGSKR 428
DB 362 LGLSAQLGIEVVMROVFFGAGNYHLVDENFPLPDYWLSTLFFKLVGTVMASVOGSKR 421
QY 429 RKLRYVLIHCTNTDNPRYKEGDITLYALINLHNTVKYLRPYPSNKQVQDKYLLRPLGPHGL 488
DB 422 SKLRVYLHCTNTDNPRYKEGDITLYALINLHNTVKYLRPYPSNKQVQDKYLLRPLGPHGL 481
QY 489 LKRSVQNLGLTLKMPDDOTLPLMEKPLRPGSSLGPAFSYFVIRNAKVAACI 543
DB 482 LKRSVQNLGLTLKMPDDOTLPLMEKPLRPGSSLGPAFSYFVIRNAKVAACI 536
RESULT 7
Q90ZF8 PRELIMINARY; PRT; 536 AA.
ID Q90ZF8
AC Q90ZF8
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Heparanase.
DE Name=Hep;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22194309; PubMed=12077130; DOI=10.1074/jbc.M203282200;
RA Podyma-Inoue K.A., Yokote H., Sakaguchi K., Ikuta M., Yanagishita M.,
RT "Characterization of heparanase from a rat parathyroid cell line."
RL J. Biol. Chem. 277.32459-32465 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Podyma K.A., Yokote H., Sakaguchi K., Ikuta M., Yanagishita M.,
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.

DR EMBL: AF184967; AA04563.1; -
DR InterPro: IPR005199; Glyco_hydro_79N.
DR Pfam: PF03662; Glyco_hydro_79n; 1.
DR SEQUENCE 536 AA; 60568 MW; 6208B1PD9EE28421 CRC64;
Query Match 74.7%; Score 2123; DB 2; Length 536;
Best Local Similarity 75.7%; Pred. No. 9.5e-151;
Matches 405; Conservative 51; Mismatches 79; Indels 0; Gaps 0;
QY 9 LPPMLLLGLPGLSPGALPPPAQADVVLDLFFTOEPHLVSPSLVTIDANLATD 68
DB 2 LRPILLMLMKRLALITGTPAGTAHPKDVVDLEFYTKRLFGQSPSLSTIDASLATD 61
QY 69 PRFLILGSPKRLTLARGSPAYLRFPGTKTDFLIFDPKKESTFEERSYWGQVQNDICK 128
DB 62 PRFLTLGSPRLRALARGSPAYLRFPGTKTDFLIFDPKKEPTSEERSYWGQVQNDICG 121
QY 129 YGSIPTDVEKRLLEMPYQOQLLRHNYQKKFKNSTYSSSDVLYTFANCGDLIFGL 188
DB 122 SERVSADVLRKLOMWPQOELLRLREOYQREPKNSTYSSSDVLYTFANCGDLIFGL 181
QY 189 NALLPTADLQNNSSNAQOLLIDYCSKGYNISWELGNEPNSFLKADIFINGSOUGEYIQ 248
DB 182 NALLPTADLQNNSSNAQOLLIDYCSKGYNISWELGNEPNSFLKADIFINGSOUGEYIQ 241
QY 249 LHKLLRKSTFKNAKLYGPDVQPRRKTAQMLKSLFKAGEVIDSVTWHYYLNGRTATRE 308
DB 242 LHKLLQKSAFQNAKLYGPDIGQPRGKTIVKLRSFLKAGEVIDSLTWHYYLNGRTATRE 301
QY 309 DFLNPDVLDLFISSVQKLYQVYESTRRPKKVMLGFTSSAYGGAPLSDTPAAGFMWLDK 368
DB 302 DFLSSVDVLDLFISSVQKLYQVYESTRRPKKVMLGFTSSAYGGAPLSDTPAAGFMWLDK 361
QY 369 LGLSARMGIEVVMROVFFGAGNYHLVDENFPLPDYWLSTLFFKLVGTVMASVOGSKR 428
DB 362 LGLSAQLGIEVVMROVFFGAGNYHLVDENFPLPDYWLSTLFFKLVGTVMASVOGSKR 421
QY 429 RKLRYVLIHCTNTDNPRYKEGDITLYALINLHNTVKYLRPYPSNKQVQDKYLLRPLGPHGL 488
DB 422 SKLRVYLHCTNTDNPRYKEGDITLYALINLHNTVKYLRPYPSNKQVQDKYLLRPLGPHGL 481
QY 489 LKRSVQNLGLTLKMPDDOTLPLMEKPLRPGSSLGPAFSYFVIRNAKVAACI 543
DB 482 LKRSVQNLGLTLKMPDDOTLPLMEKPLRPGSSLGPAFSYFVIRNAKVAACI 536
RESULT 8
Q90YK5 PRELIMINARY; PRT; 523 AA.
ID Q90YK5
AC Q90YK5
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Heparanase.
DE Gallus gallus (Chicken).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21369599; PubMed=1187326; DOI=10.1074/jbc.M102462200;
RA Goldsmith O., Zcharia E., Aingorn H., Guatta-Rangani Z., Atzmon R.,
RA Michael I., Becker I., Mitrani E., Vlodavsky I.,
RT "Expression pattern and secretion of human and chicken heparanase are
determined by their signal peptide sequence."
RL J. Biol. Chem. 276.22178-22187 (2001).
DR EMBL: AY037007; AAK82648.1; -
DR Pfam: PF03662; Glyco_hydro_79n; 1.
DR SEQUENCE 523 AA; 58386 MW; 8EB0B7B18C9BF881 CRC64;
Query Match 57.9%; Score 1645.5; DB 2; Length 523;
Best Local Similarity 60.2%; Pred. No. 6.8e-115;

RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takehashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
 RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN
 RN SEQUENCE FROM N.A.
 RC STRAIN=JUL/J; TISSUE=Spleen;
 RA Huilett M.D., Wang J., Hornby J.R., Freeman C., Pagler E., McHenry J.,
 RA Parish C.R.,
 RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL; A077467; AAL76083.1; -
 DR EMBL; AK040471; BAC30600.1; -
 DR EMBL; AF359507; AA015188.1; -
 DR MGD; MG1:134324; Hspc.
 DR GO; GO:0005578; C:extracellular matrix (sensus Metazoa); TAS.
 DR InterPro; IPR005199; Glyco_hydro_79n; I.
 DR Pfam; PF03662; Glyco_hydro_79n; I.
 SQ SEQUENCE 535 AA; 60065 MW; 6E7A8302FB8A0DF CRC64;

Query Match 75.7%; Score 2150; DB 2; Length 535;
 Best Local Similarity 76.6%; Pred. No. 8.9e-153;
 Matches 407; Conservative 51; Mismatches 73; Indels 0; Gaps 0;

QY 13 LMILLGRLPLSPGALPPRAQADVVDLDFTOEPLHVSFSLSVITDANLADPRL 72
 DB 5 LLLMLWGPGLAQAQAPACTAPTDVVDLEFYTKRLRSVSPFLSITIDASLATDPR 64
 QY 73 ILGSPKRLTLARGSPAYLRFGCTKTDPLDPKKESTFEERSYQWQOVNDICRYGSI 132
 DB 65 TFLGSPRLALARGSPAYLRFGCTKTDPLDPKKESTFEERSYQWQOVNDICRYGSI 124
 QY 133 PDVEEKLRLQVEMPOQLLREHYOKKFNKSTYSRSSVDVLYTFANCGLDLIFGNALL 192
 DB 125 SAAVLRKLQVEMPOQLLREHYOKKFNKSTYSRSSVDVLYTFANCGLDLIFGNALL 184
 QY 193 RTADLQWSSNAQLLDYSSKGYNISWELGNEPNSFLKKADI FINGSLGEDIYOLHKL 252
 DB 185 RRPDLRMWSSNAQLLDYSSKGYNISWELGNEPNSFWKKAHLIDGLDGFVFLHKL 244
 QY 253 LRKSTFKNAKLYGPPVGOVRRRTAKMLKSFKAAGEVIDSVTHHHYLLNGRTATREDPLN 312
 DB 245 LQSAFQNAKLYGPPVGOVRRRTAKMLKSFKAAGEVIDSVTHHHYLLNGRTATREDPLN 304
 QY 313 PDVLIFFISSVQKVPQVVESTPGKKVWLGKTSAYAGGAPLLSDTFAAGFMWLDKLGIS 372
 DB 305 SDVLDLFFISSVQKILKVTKEITPGKKVWLGKTSAYAGGAPLLSDTFAAGFMWLDKLGIS 364
 QY 373 ARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLFFKLVGTKYLMASVQGSKRKRLR 432
 DB 365 AQMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLFFKLVGTKYLMASVQGSKRKRLR 424
 QY 433 VYLHCTNTDNPYKKGDLTLVAINLHNVTKYRLPYPSFNKOVDKYLRPLGPHGLSKS 492
 DB 425 VYLHCTNVHNPYKKGDLTLVAINLHNVTKYRLPYPSFNKOVDKYLRPLGPHGLSKS 484
 QY 493 VOLNGILTKMVDQTLPLMEKPLRPGSSGLPAPSYGFVIRNAKVAACI 543
 DB 485 VOLNGILTKMVDQTLPLMEKPLRPGSSGLPAPSYGFVIRNAKVAACI 535

RESULT 5
 O6YGZ1 PRELIMINARY; PRT; 535 AA.
 AC O6YGZ1; 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DE Hepatanae.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
 RN
 RN SEQUENCE FROM N.A.
 RP MEDLINE=22841152; PubMed=12837765; DOI=10.1074/jbc.M300925200;
 RX Gong F., Uemth P., Galvis M.L.E., Vlodevsky I., Horner A., Lindahl U.,
 RA Li J.P.,
 RA "Processing of macromolecular heparin by heparanase."
 RT
 RL J. Biol. Chem. 278:35152-35158(2003).
 RN
 RN SEQUENCE FROM N.A.
 RA Li J.-P., Gong F., Lindahl U.,
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY151051; AAN41636.1; -
 DR GO; GO:0005578; C:extracellular matrix (sensus Metazoa); TAS.
 DR InterPro; IPR005199; Glyco_hydro_79n; I.
 DR Pfam; PF03662; Glyco_hydro_79n; I.
 SQ SEQUENCE 535 AA; 59992 MW; 3748ABE3795C718A CRC64;

Query Match 75.3%; Score 2140; DB 2; Length 535;
 Best Local Similarity 76.6%; Pred. No. 5e-152;
 Matches 407; Conservative 50; Mismatches 74; Indels 0; Gaps 0;

QY 13 LMILLGRLPLSPGALPPRAQADVVDLDFTOEPLHVSFSLSVITDANLADPRL 72
 DB 5 LLLMLWGPGLAQAQAPACTAPTDVVDLEFYTKRLRSVSPFLSITIDASLATDPR 64
 QY 73 ILGSPKRLTLARGSPAYLRFGCTKTDPLDPKKESTFEERSYQWQOVNDICRYGSI 132
 DB 65 TFLGSPRLALARGSPAYLRFGCTKTDPLDPKKESTFEERSYQWQOVNDICRYGSI 124
 QY 133 PDVEEKLRLQVEMPOQLLREHYOKKFNKSTYSRSSVDVLYTFANCGLDLIFGNALL 192
 DB 125 SAAVLRKLQVEMPOQLLREHYOKKFNKSTYSRSSVDVLYTFANCGLDLIFGNALL 184
 QY 193 RTADLQWSSNAQLLDYSSKGYNISWELGNEPNSFLKKADI FINGSLGEDIYOLHKL 252
 DB 185 RRPDLRMWSSNAQLLDYSSKGYNISWELGNEPNSFWKKAHLIDGLDGFVFLHKL 244
 QY 253 LRKSTFKNAKLYGPPVGOVRRRTAKMLKSFKAAGEVIDSVTHHHYLLNGRTATREDPLN 312
 DB 245 LQSAFQNAKLYGPPVGOVRRRTAKMLKSFKAAGEVIDSVTHHHYLLNGRTATREDPLN 304
 QY 313 PDVLIFFISSVQKVPQVVESTPGKKVWLGKTSAYAGGAPLLSDTFAAGFMWLDKLGIS 372
 DB 305 SDVLDLFFISSVQKILKVTKEITPGKKVWLGKTSAYAGGAPLLSDTFAAGFMWLDKLGIS 364
 QY 373 ARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLFFKLVGTKYLMASVQGSKRKRLR 432
 DB 365 AQMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLFFKLVGTKYLMASVQGSKRKRLR 424
 QY 433 VYLHCTNTDNPYKKGDLTLVAINLHNVTKYRLPYPSFNKOVDKYLRPLGPHGLSKS 492
 DB 425 VYLHCTNVHNPYKKGDLTLVAINLHNVTKYRLPYPSFNKOVDKYLRPLGPHGLSKS 484
 QY 493 VOLNGILTKMVDQTLPLMEKPLRPGSSGLPAPSYGFVIRNAKVAACI 543
 DB 485 VOLNGILTKMVDQTLPLMEKPLRPGSSGLPAPSYGFVIRNAKVAACI 535

RESULT 6
 O7IRP1 PRELIMINARY; PRT; 536 AA.
 AC O7IRP1; 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DE Hepatanae.
 GN Name=Hspc;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]

```

RESULT 3
ID 09MYO PRELIMINARY; PRT; 545 AA.
AC 09MYO
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Heparanase.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=21176669; PubMed=11277877;
RA Kizaki K., Nakano H., Nakano H., Takahashi T., Imai K., Hashizume K.;
RT "Expression of heparanase mRNA in bovine placenta during gestation.";
RL Reproduction 121:573-580(2001).
DR EMBL, AF281160; AAF87301.2;
DR InterPro; IPR005199; Glyco_Hydro_79N.
DR Pfam; PF03662; Glyco_Hydro_79n; 1.
SQ SEQUENCE 545 AA; 61076 MW; FAC4BDFD85B933 CRC64;

Query Match 80.3%; Score 2282; DB 2; Length 545;
Best Local Similarity 79.8%; Pred. No. 1.1e-162;
Matches 435; Conservative 35; Mismatches 73; Indels 2; Gaps 1;

QY 1 MLARKPALPPLMLL--LGLPLGLPLPGALPRPAQADVDLDLPTQLHLVSPFLS 58
DB 1 MLACRKPGLRPLPLLLPLGLPGSPCTPAAPADDAELEPFTERRPLHLVSPAFLS 60
QY 59 VTIDANLATDPRPLILLSGPKRTIARGLSPAYLRGCKTKTPIIFDKKESTFEERSY 118
DB 61 FTIDANLATDPRFFFLGSSKRTIARGLAPYLRGKNGKGFLLPDKKEPAFEERSY 120
QY 119 OSQVNDICKYSGIPDVEEKLRLWPYQEOQLLRHYOKKFNSTYSRSSVDVLYTFAN 178
DB 121 LSQSNODICKSGSPDVEEKLRLWPYQEOQLLRHYOKKFNSTYSRSSVDVLYTFAS 180
QY 179 CSGLDLIFGLNALRLTADLQWNSNAQLLDYCSSKNGYISWELGNEPNSFLKADIFIN 238
DB 181 GSGMLNIFGVNMLRLRTDMHMSNAQLLDYCSSKNGYISWELGNEPNSFLKADIFIN 240
QY 239 GSGLDGDIYQLHKLKSTFKPAKLYGPDVGGPRKRTAKMLSGFLKAGGEVIDSVTWHY 298
DB 241 GSGLDGDIYQLHKLKSTFKPAKLYGPDVGGPRKRTAKMLSGFLKAGGEVIDSVTWHY 300
QY 299 YNGRATATEDPLNPVDLIFITSSVQKFOVVESTRPGKKVMGERTSSAYGGAPLSDT 358
DB 301 YNGRATATEDPLNPVDLIFITSSVQKFOVVESTRPGKKVMGERTSSAYGGAPLSDT 360
QY 359 PAAGFWMLDKLGLSARMGIEVWNRQVFGAGNYHLVDENFDPLPDYMLLFFKLVGY 418
DB 361 PAAGFWMLDKLGLSARMGIEVWNRQVFGAGNYHLVDENFDPLPDYMLLFFKLVGY 420
QY 419 LMASVGSGRKRIKRYLHCTNTDNPYKGGDLTLVYINLHNTKYLRLPFPKNOVDY 478
DB 421 LMASVGSGRKRIKRYLHCTNTDNPYKGGDLTLVYINLHNTKYLRLPFPKNOVDY 480
QY 479 LRLPGLPHGLKSVYQNLGLTKMVDQTLPLMEKPLRPGSSILGPAFYSFVIRNKK 538
DB 481 LRLPGLPHGLKSVYQNLGLTKMVDQTLPLMEKPLRPGSSILGPAFYSFVIRNKK 540
QY 539 VAACI 543
DB 541 VAACI 545

```

```

AC 08K3K3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Heparanase (Mus musculus 0 day neonate thymus cDNA, RIKEN full-length
DE enriched library, clone:A43010M04 product:heparanase, full insert
DE sequence).
GN Name=Hpe; Synonyms=Hpa, Hspe;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB;
RA Miao H.-Q., Navarro B., Patel S., Sargent D., Koo H., Wan H.,
RA Plata A., Zhou Q., Ludwig D., Bohlen P., Kussie P.;
RT "Cloning, expression, and purification of mouse heparanase.";
RL Protein Expr. Purif. 0:0-0(2002).
(2)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
(3)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
(4)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60 770 full-length cDNAs.";
RL Nature 420:563-573(2002).
(5)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Komo H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
(6)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20530913; PubMed=11076661; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Saeki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuunai T., Taishiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujitake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multitepillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
(7)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai U., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,

```

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schermer A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RL (5)
 RP SEQUENCE FROM N.A.
 RC Tissue=pancreas;
 RN Strausberg R.;
 RL Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Vlodavsky I., Friedman Y., Elkin M., Aingorn H., Arzon R.,
 RA Ishai-Michaeli R., Bitan M., Pappo O., Peretz T., Michael I.,
 RA Spector L., Becker I.;
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF161514; AAD45379.1; -;
 DR EMBL; AF152376; AAD45669.1; -;
 DR EMBL; AF15510; AAD54941.1; -;
 DR EMBL; BC051321; AAN51321.1; -;
 DR EMBL; AF144325; AAD41342.1; -;
 DR GO; GO:0004566; F-beta-glucuronidase activity; TAS.
 DR GO; GO:0006029; P-glycoprotein-mediated transport; TAS.
 DR InterPro; IPR005199; Glyco_hydro_79N.
 DR Pfam; PF03662; Glyco_hydro_79n; 1.
 DR CHAIN 158 543 heparanase.
 SQ SEQUENCE 543 AA; 61176 MW; AD262EC267334AB2 CRC64;

Query Match 99.9%; Score 2838; DB 2; Length 543;
 Best Local Similarity 99.8%; Pred. No. 1.9e-204;
 Matches 542; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLRSKPALPPMLLLGLPLSGALPPRQAQDVVDLDFPQEPHLVSPSLSTT 60
 DB 1 MLRSKPALPPMLLLGLPLSGALPPRQAQDVVDLDFPQEPHLVSPSLSTT 60
 QY 61 IDANLATDPRFLLILGSPKRLTARGLSPAYLRFSGTKTDPLFPDKESTFEERSY 120
 DB 61 IDANLATDPRFLLILGSPKRLTARGLSPAYLRFSGTKTDPLFPDKESTFEERSY 120
 QY 121 QVNODICKYGSIPPDVEEKLRLMPYQEOULLREHYOKKFKNSTYSSVDVLYTFAN 180
 DB 121 QVNODICKYGSIPPDVEEKLRLMPYQEOULLREHYOKKFKNSTYSSVDVLYTFAN 180
 QY 181 GDLDFGLNALRTADLQWSSNAQLLDYCSSKGVNISMELGNENPSFLKADIFIN 240
 DB 181 GDLDFGLNALRTADLQWSSNAQLLDYCSSKGVNISMELGNENPSFLKADIFIN 240
 QY 241 QGGEVDYIQHLKLRKSTFFKNAKLYGPDVGPQRRTAKMLKSPKAGGEVIDSVTHHY 300
 DB 241 QGGEVDYIQHLKLRKSTFFKNAKLYGPDVGPQRRTAKMLKSPKAGGEVIDSVTHHY 300
 QY 301 NGRTATREDPLNDVDLDFISSVQKFFQVVESTRPGKWLGETSSAYGGAPLSDT 360
 DB 301 NGRTATREDPLNDVDLDFISSVQKFFQVVESTRPGKWLGETSSAYGGAPLSDT 360
 QY 361 AGFMWLDKLGSLARMGIEVVMROVFFGAGNYHLVDENFDPDLDPYWSLFLFKLVGT 420
 DB 361 AGFMWLDKLGSLARMGIEVVMROVFFGAGNYHLVDENFDPDLDPYWSLFLFKLVGT 420
 QY 421 ASVQSGSKRRKRLRYVLTNTDNPRIYEGDLTLVAINLHNTKTLRLPYPSNKOVD 480
 DB 421 ASVQSGSKRRKRLRYVLTNTDNPRIYEGDLTLVAINLHNTKTLRLPYPSNKOVD 480
 QY 481 RPLGPHGLSKSVQNLGLTLKAVDDQTLPLMEKPLRPGSSLGLPASYSFPIYIRNA 540
 DB 481 RPLGPHGLSKSVQNLGLTLKAVDDQTLPLMEKPLRPGSSLGLPASYSFPIYIRNA 540
 QY 541 ACI 543
 DB 541 ACI 543

RESULT 2
 ID Q9UL39 PRELIMINARY; PRT; 545 AA.
 AC Q9UL39;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Heparanase.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Placenta;
 RX MEDLINE=20229546; PubMed=10764835; DOI=10.1093/glycob/10.5.467;
 RA Dempsey L.A., Plummer T.B., Coombes S.L., Platt J.L.;
 RT "Heparanase expression in invasive trophoblasts and acute vascular
 RT damage.";
 RL Glycobiology 10:467-475(2000).
 DR EMBL; AF084467; AAD54516.1; -;
 DR Genew; HGNC:5164; HSE.
 DR InterPro; IPR005199; Glyco_hydro_79N.
 DR Pfam; PF03662; Glyco_hydro_79n; 1.
 SQ SEQUENCE 545 AA; 61417 MW; 67B80ACD735A9A1 CRC64;

Query Match 99.1%; Score 2817; DB 2; Length 545;
 Best Local Similarity 99.4%; Pred. No. 7.1e-203;
 Matches 542; Conservative 1; Mismatches 0; Indels 2; Gaps 2;

QY 1 MLRSKPALPPMLLLGLPLSGALPPRQAQDVVDLDFPQEPHLVSPSLSTT 58
 DB 1 MLRSKPALPPMLLLGLPLSGALPPRQAQDVVDLDFPQEPHLVSPSLSTT 60
 QY 59 VTIDANLATDPRFLLILGSPKRLTARGLSPAYLRFSGTKTDPLFPDKESTFEERSY 118
 DB 61 VTIDANLATDPRFLLILGSPKRLTARGLSPAYLRFSGTKTDPLFPDKESTFEERSY 120
 QY 119 QVNODICKYGSIPPDVEEKLRLMPYQEOULLREHYOKKFKNSTYSSVDVLYTFAN 178
 DB 121 QVNODICKYGSIPPDVEEKLRLMPYQEOULLREHYOKKFKNSTYSSVDVLYTFAN 180
 QY 179 GDLDFGLNALRTADLQWSSNAQLLDYCSSKGVNISMELGNENPSFLKADIFIN 238
 DB 181 GDLDFGLNALRTADLQWSSNAQLLDYCSSKGVNISMELGNENPSFLKADIFIN 240
 QY 238 GDLDFGLNALRTADLQWSSNAQLLDYCSSKGVNISMELGNENPSFLKADIFIN 298
 DB 241 GDLDFGLNALRTADLQWSSNAQLLDYCSSKGVNISMELGNENPSFLKADIFIN 300
 QY 299 YNGRTATREDPLNDVDLDFISSVQKFFQVVESTRPGKWLGETSSAYGGAPLSDT 358
 DB 301 YNGRTATREDPLNDVDLDFISSVQKFFQVVESTRPGKWLGETSSAYGGAPLSDT 360
 QY 359 PAAGFMWLDKLGSLARMGIEVVMROVFFGAGNYHLVDENFDPDLDPYWSLFLFKLVGT 418
 DB 361 PAAGFMWLDKLGSLARMGIEVVMROVFFGAGNYHLVDENFDPDLDPYWSLFLFKLVGT 420
 QY 419 LMASVQSGSKRRKRLRYVLTNTDNPRIYEGDLTLVAINLHNTKTLRLPYPSNKOVD 478
 DB 421 LMASVQSGSKRRKRLRYVLTNTDNPRIYEGDLTLVAINLHNTKTLRLPYPSNKOVD 480
 QY 479 LRPGLGPHGLSKSVQNLGLTLKAVDDQTLPLMEKPLRPGSSLGLPASYSFPIYIRNA 538
 DB 481 LRPGLGPHGLSKSVQNLGLTLKAVDDQTLPLMEKPLRPGSSLGLPASYSFPIYIRNA 540
 QY 539 VAACT 543
 DB 541 VAACT 545

091159207

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2005, 11:18:07 ; Search time 109.5 Seconds
(without alignments)
2539.352 Million cell updates/sec

Title: SEQ2A
Perfect score: 2842
Sequence: 1 MLRSKPLPPLMLLGP.....LPFSYSPFVIRNAVAACI 543

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2838	99.9	543	2	Q9Y251 homo sapien
2	2817	99.1	545	2	Q9UJ39 homo sapien
3	2282	80.3	545	2	Q9MYX0 Bos taurus
4	2150	75.7	535	2	Q8K3K3 mus musculu
5	2140	75.3	535	2	Q6VGT1 mus musculu
6	2135	75.1	536	2	Q71RPI ratu
7	2133	74.7	536	2	Q9QZF8 ratu
8	1645.5	57.9	523	2	Q90YK5 gallus gall
9	1154.5	40.6	592	2	Q9HB37 homo sapien
10	1146.5	40.3	592	2	Q8WMQ2 homo sapien
11	1015.5	35.7	548	2	Q8WMQ1 homo sapien
12	936.5	33.0	534	2	Q9HB38 homo sapien
13	897.5	31.6	480	2	Q9HB39 homo sapien
14	696	24.5	515	2	Q8T108 bombyx mori
15	416	14.6	521	2	Q9SDA1 aradidopsi
16	416	14.6	543	2	Q9FF10 aradidopsi
17	404	14.2	559	2	Q89F99 bradyrhizob
18	392.5	13.8	544	2	Q8H615 oryza sativ
19	381	13.4	527	2	Q91T08 scutellaria
20	379	13.3	541	2	Q69J16 oryza sativ
21	364	12.8	517	2	Q70YJ3 hordea vul
22	363	12.8	536	2	Q9FZP1 aradidopsi
23	352.5	12.4	516	2	Q9F1K8 aradidopsi
24	352.5	12.4	539	2	Q8L608 aradidopsi
25	350.5	12.3	529	2	Q6Z0E2 oryza sativ
26	169.5	6.0	190	2	Q82604 aradidopsi
27	160	5.6	935	2	Q9VE79 drosophila
28	141	5.0	559	2	Q7SFB0 neurospora
29	136.5	4.8	463	2	Q63T97 burkholderi
30	130.5	4.6	493	2	Q9HK01 thermoplasma
31	122.5	4.3	408	2	Q9HEZ1 phanerochaete

32	122.5	4.3	408	2	Q9HEZ2 phanerochaete
33	116.5	4.1	617	2	Q40996 measles vir
34	114.5	4.0	1169	2	Q869K5 dictyostell
35	114	4.0	398	2	Q72RP7 leptospira
36	114	4.0	398	2	Q8F410 leptospira
37	113.5	4.0	617	2	Q83295 measles vir
38	113.5	4.0	1829	2	Q9KH44 parvovirus
39	113	4.0	390	2	Q8T9H7
40	112.5	4.0	356	1	Y670 METUA
41	112.5	4.0	617	2	Q83647
42	112	3.9	732	2	Q7MSD4
43	111.5	3.9	493	2	Q979W0
44	111.5	3.9	575	2	Q43855
45	111.5	3.9	617	2	Q40991 measles vir

ALIGNMENTS

RESULT 1	ID	Q9Y251	PRELIMINARY;	PRT;	543 AA.
AC	Q9Y251	01-NOV-1999 (TrEMBLrel. 12, Created)			
DT	01-NOV-1999 (TrEMBLrel. 12, Last sequence update)				
DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)				
DE	Heparanase.				
GN	Name=HPA; Synonyms=HPSE;				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RP	SEQUENCE FROM N.A.				
RC	TISUE=Placenta;				
RX	MEDLINE=99321249; PubMed=10395326;				
RA	Hulec M.D., Freeman C., Hamdorf B.J., Baker R.T., Harris M.J.,				
RA	Parish C.R.;				
RT	"Cloning of mammalian heparanase, an important enzyme in tumor				
RT	invasion and metastasis";				
RL	Nat. Med. 5:803-809(1999).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISUE=Placenta;				
RX	MEDLINE=99335379; PubMed=10405343; DOI=10.1006/brc.1999.0962;				
RA	Kussie P.H., Holmes J.D., Ludwig D.L., Patel S., Navarro E.C.,				
RA	Seddon A.P., Giorgio N.A., Bohlen P.;				
RT	"Cloning and functional expression of a human heparanase gene";				
RT	Biochem. Biophys. Res. Commun. 261:183-187(1999).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99377052; PubMed=10446189; DOI=10.1074/jbc.274.34.24153;				
RA	Toyoshima M., Nakajima M.;				
RT	"Human heparanase. Purification, characterization, cloning, and				
RT	expression";				
RL	J. Biol. Chem. 274:24153-24160(1999).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	TISUE=Pancreas;				
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242601899;				
RA	Strausberg R.L., Feigold E.A., Grouse L.H., Derge J.G.,				
RA	Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,				
RA	Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,				
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Scapleton M., Soares M.B., Bonaldo W.F., Casavant T.L., Scheetz T.E.,				
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Pange C.,				
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Millan S.J.,				
RA	Bosch S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,				
RA	Villalon D.K., Murny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA	Faney J., Heiton E., Kettelman W., Madan A., Rodriguez S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)